

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2003, 20:07:47 ; Search time 42 Seconds  
(without alignments)  
1067.014 Million cell updates/sec

Title: US-09-996-223-1

Sequence: 1 MALTITAAARLGTNASCCLV.....ERPKSMSTEGIMKFDVSKSG 466

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: PIR.76.\*
- 2: PIR1.\*
- 3: PIR2.\*
- 4: PIR3.\*
- 5: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2339	96.0	464	1	YKPG citrate (sl)-synth
2	1693	69.5	468	2	S42370 citrate (sl)-synth
3	1519	62.4	473	2	T39028 citrate synthase P
4	1501.5	61.6	469	2	T49379 citrate synthase,
5	1499.5	61.6	479	1	YKBY citrate (sl)-synth
6	1498.5	61.5	474	2	T02390 citrate (sl)-synth
7	1495.5	61.4	469	2	S41563 citrate (sl)-synth
8	1473.5	60.5	433	2	T47834 citrate synthase-1
9	1425	58.5	460	1	YKBYC citrate (sl)-synth
10	1349	55.4	471	2	S44316 citrate (sl)-synth
11	1219	50.0	462	2	JC5625 14-nm filament pro
12	1157.5	47.5	472	1	YKQDM citrate (sl)-synth
13	1108.5	45.5	339	2	T09334 citrate (sl)-synth
14	1063.5	43.7	486	2	S52814 citrate (sl)-synth
15	356	14.6	437	2	I40717 citrate (sl)-synth
16	351	14.4	436	1	YKOPC citrate (sl)-synth
17	346	14.2	437	2	S74344 citrate synthase 9
18	345.5	14.2	427	2	B81139 citrate synthase-1
19	339.5	13.9	514	2	T49158 citrate synthase-1
20	338	13.9	428	1	YKPSA citrate (sl)-synth
21	338	13.9	428	2	B84448 citrate synthase P
22	338	13.9	509	2	C84858 citrate synthase s
23	337.5	13.6	431	2	E70782 citrate synthase 1
24	331.5	13.6	431	2	T45353 citrate synthase 1
25	327.5	13.4	430	2	J01392 citrate (sl)-synth
26	327.5	13.4	435	2	D97863 citrate (sl)-synth
27	324.5	13.3	411	2	A75479 citrate synthase -
28	324	13.3	426	2	E87485 citrate synthase 1
29	322.5	13.2	372	2	H84044 citrate synthase 1

30	322.5	13.2	516	2	S53007 citrate synthase -
31	320	13.1	441	2	F86708 citrate (sl)-synth
32	318	13.1	378	2	AF1834 citrate synthase 1
33	315.5	13.0	427	1	YKRC citrate (sl)-synth
34	315.5	13.0	427	2	A93722 citrate synthase 1
35	315	12.9	437	2	AF3356 citrate (sl)-synth
36	314	12.9	424	2	I39506 citrate synthase X
37	313	12.8	429	2	D82669 citrate synthase 1
38	311.5	12.8	436	1	YKRECP citrate (sl)-synth
39	309.5	12.7	427	2	H85572 citrate synthase 1
40	309	12.7	431	2	B82119 citrate synthase Y
41	307.5	12.6	431	2	I40044 citrate (sl)-synth
42	304	12.5	429	2	AH2747 citrate synthase 1
43	304	12.5	429	2	G97528 citrate synthase 1
44	303.5	12.5	367	2	B72394 citrate synthase -
45	303.5	12.5	427	2	AF0590 citrate synthase 1

## ALIGNMENTS

### RESULT 1

YKPG citrate (sl)-synthase (EC 4.1.3.7) precursor - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 15-Oct-1982 #sequence,revision 30-Sep-1992 #text\_change 05-May-2000

C:Accession: A29966; A01109; A61347

R:Evans, C.T.; Owens, D.D.; Sumegi, B.; Kispal, G.; Stere, P.A.

A:Title: Isolation, nucleotide sequence, and expression of a cDNA encoding pig citr

A:Reference number: A29966; MUID:8900065; PMID:3048387

A:Accession: A29966

A:Molecule type: mRNA

A:Residues: 1-464 <EY>

A:Cross-references: EMBL:M21197; NID:G164418; PIDN:AA31017.1; PID:G164419

R:Blotnam, D.P.; Parmelee, D.C.; Kumar, S.; Walsh, K.A.; Titani, K.

R:Biochemstry 21, 2028-2036, 1982

A:Title: Complete amino acid sequence of porcine heart citrate synthase:

A:Reference number: A90457; MUID:82231993; PMID:7093227

A:Accession: A01109

A:Molecule type: Protein

A:Residues: 28-464 <BL>

R:Blotnam, D.P.; Parmelee, D.C.; Kumar, S.; Wade, R.D.; Ericsson, L.H.; Neurath, H.;

Proc. Natl. Acad. Sci. U.S.A. 78, 5381-5385, 1981

A:Title: Primary structure of porcine heart citrate synthase.

A:Reference number: A61347; MUID:82060250; PMID:6795632

A:Accession: A61347

A:Molecule type: Protein

A:Residues: 28-464 <BL>

R:Remington, S.; Wiegand, G.; Huber, R.

J. Mol. Biol. 158, 111-152, 1982

A:Title: Crystallographic refinement and atomic models of two different forms of cit

A:Reference number: A92884; MUID:83010291; PMID:7120407

A:Contents: annotation; X-ray crystallography, 2.7 and 1.7 angstroms

C:Comment: Citrate (sl)-synthase is found in nearly all cells capable of oxidative i

ycle.

C:Comment: It is synthesized in the cytoplasm but functions in the mitochondrion of

C:Comment: This molecule is a dimer of identical chains. Each dimer binds two molecu

C:Superfamily: citrate (sl)-synthase

C:Keywords: carbon-carbon lyase; coenzyme A; homodimer; methylated amino acid; mitoc

F:1-27/Domains: transit peptide (mitochondrion) #status predicted <Sig>

F:28-464/Product: citrate (sl)-synthase #status experimental <MNT>

F:301,347,402/Active site: His, His, Asp #status predicted

F:395/Modified site: N6,N6,N6-trimethyllysine (lys) #status experimental

Query Match 96.0%; Score 2339; DB 1; Length 464;

Best Local Similarity 96.1%; Pred. No. 5.8e-169;

Matches 446; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

OY 1 MALTITAAARLGTNASCCLVLAARHSAASSTNKLADIIPKROARKTFROOHGKTVV 60  
DB 1 MALTITAAARLGTNASCCLVLAARHSAASSTNKLADIIPKROARKTFROOHGKTVV 60

**THIS PAGE BLANK (USPTO)**

QY 61 GQITVDMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPEGLF 120  
 |||  
 DB 61 GQITVDMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPEGLF 120  
 QY 121 WILVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNSBN 180  
 |||  
 DB 121 WILVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNSBN 180  
 QY 181 FARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSHNF 240  
 |||  
 DB 181 FARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSHNF 240  
 QY 241 TMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNSBN 300  
 |||  
 DB 241 TMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNSBN 300  
 QY 301 HCLANOEVVWLTOLQKEVGVDSDEKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYTQ 360  
 |||  
 DB 301 HCLANOEVVWLTOLQKEVGVDSDEKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYTQ 360  
 QY 361 REFALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 420  
 |||  
 DB 361 REFALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 420  
 QY 421 TTVLFGVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 464  
 |||  
 DB 421 TTVLFGVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 464

## RESULT 2

citrate (si)-synthase (EC 4.1.3.7) precursor - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 07-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 05-May-2000  
 C:Accession: S42370  
 R:Smith, A.  
 Submitted to the EMBL Data Library, March 1994  
 A:Reference number: S42368  
 A:Accession: S42370  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-468 <SMI>  
 A:Cross-references: EMBL:Z30423; NID:G458479; PID:G458482  
 C:Genetics: 69/3; 202/3; 309/3  
 A:Introns: citrate (si)-synthase  
 C:Superfamily: citrate (si)-synthase  
 C:Keywords: carbon-carbon lyase; coenzyme A; mitochondrion; oxo-acid-lyase

Query Match 69.5%; Score 1693; DB 2; Length 468;

Best Local Similarity 68.3%; Pred. No. 3,9e-120; Matches 315; Conservative 62; Mismatches 82; Indels 2; Gaps 1;

QY 1 MALLTAARLLGKRNA--SCLVLAARHASSTNLKDIADLPKEDARIKTRROHGK 58  
 |||  
 DB 1 MSLSGMARIRLITKGVIPQVAPLSTSAEGSTLKEVLSKIPAHNAKVSRTERTG 60  
 QY 59 VNGQITVDMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPEG 118  
 |||  
 DB 59 VNGQITVDMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPEG 118  
 QY 119 LFWLLVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 178  
 |||  
 DB 119 LFWLLVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 178  
 QY 121 IWWLLCTGDPVSEQTAITIKEMARADLPHTVYRMLDNPTNLHPMAQFIAATAALNNE 180  
 |||  
 DB 121 IWWLLCTGDPVSEQTAITIKEMARADLPHTVYRMLDNPTNLHPMAQFIAATAALNNE 180  
 QY 179 SNEFARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSH 238  
 |||  
 DB 179 SNEFARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSH 238  
 QY 239 NFWTMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 298  
 |||  
 DB 239 NFWTMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 298  
 QY 241 NFWTMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 300  
 |||  
 DB 241 NFWTMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 300  
 QY 299 PLHGLANOEVVWLTOLQKEVGVDSDEKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYT 358

DB 301 PLHGLANOEVVWLTOLQKEVGVDSDEKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYT 360  
 |||  
 QY 359 COREPALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 418  
 |||  
 DB 361 COREPALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 420  
 QY 419 YTVLFGVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 459  
 |||  
 DB 421 YTVLFGVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 461

## RESULT 3

citrate synthase precursor, mitochondrial - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
 C:Accession: T39028  
 R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 Submitted to the EMBL Data Library, February 1996  
 A:Reference number: Z21750  
 A:Accession: T39028  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-473 <DEV>  
 A:Cross-references: EMBL:Z69731; PID:CAA93617.2; GSPDB:GNO0066; SPDB:SPAC6C3.04  
 A:Experimental source: strain 972h-; cosmid c6C3  
 C:Genetics:  
 A:Gene: SPDB:SPAC6C3.04  
 A:Map position: 1  
 A:Genome: nuclear  
 A:Superfamily: citrate (si)-synthase  
 C:Keywords: mitochondrion

Query Match 62.4%; Score 1519; DB 2; Length 473;

Best Local Similarity 61.3%; Pred. No. 5,6e-107; Matches 287; Conservative 75; Mismatches 96; Indels 10; Gaps 3;

QY 6 AARLLGTAKNACSLVLAARHA-----SASTNLKDIADLPKEDARIKTRROHGK 57  
 |||  
 DB 7 ASTRKLASSLLSOASLRQNLNPLFTSSYSTSSSLKDLRLAIEPKQEIKKFRAHQ 66  
 QY 58 TVVGOITVDMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPE 117  
 |||  
 DB 58 TVVGOITVDMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPE 117  
 QY 67 DVIGEVTTINOMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPE 126  
 |||  
 DB 67 DVIGEVTTINOMYGGMRGKGLVYENSVLDPDEGIRFREGFSTPECOKLLPKAKGGEPLPE 126  
 QY 118 GFWLLVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 177  
 |||  
 DB 118 GFWLLVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 177  
 QY 127 SLFWLLVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 186  
 |||  
 DB 127 SLFWLLVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 186  
 QY 178 SNEFARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSH 237  
 |||  
 DB 178 SNEFARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSH 237  
 QY 238 HNTMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 296  
 |||  
 DB 238 HNTMLGVTGHIPTEREOVWMLSEKEMAKRAALPSHVVTMLDNPTNLHPMSQLSAAATALNS 296  
 QY 246 YNEFARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSH 305  
 |||  
 DB 246 YNEFARAVAGISRTKYWELIYEDSMDLIAKPCVAARIYRNLREGSGIGAIIDSLDWSH 305  
 QY 297 AGPLHGLANOEVVWLTOLQKEVGVDSDEKLDYIMNTLNSGRVVPYGYHAVALRKTDPR 356  
 |||  
 DB 297 AGPLHGLANOEVVWLTOLQKEVGVDSDEKLDYIMNTLNSGRVVPYGYHAVALRKTDPR 356  
 QY 357 YTCOREPALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 416  
 |||  
 DB 357 YTCOREPALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 416  
 QY 366 YTAOREPALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 425  
 |||  
 DB 366 YTAOREPALKHLPNDPMPFKLVAOLYKIVPNVLLDQKAKNPMPNDASHGVLLQYTGTEMY 425  
 QY 417 MNTYVLFVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 464  
 |||  
 DB 417 MNTYVLFVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 464  
 QY 426 QSEFVTVLFGVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 473  
 |||  
 DB 426 QSEFVTVLFGVSRALGVLAOLIMSRALGFPLERPKSSTEGIMKFDVSK 473

## RESULT 4

T49379

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2003, 20:02:07 ; Search time 43 Seconds  
(without alignments)  
1720.153 Million cell updates/sec

Title: US-09-996-223-1  
Perfect score: 2436  
Sequence: 1 MALLTAARLIGTKNASCLV.....ERPKSMSTEGLMKRFVDSKSG 466

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/genesep-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/genesep-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/genesep-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/genesep-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/genesep-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	100.0	466	22	AAAG3432
2	2436	100.0	466	23	ABG32120
3	2436	100.0	466	23	ABB05627
4	2339	96.0	464	23	ABG32121
5	1737.5	71.3	464	22	ABB60227
6	1498.5	61.5	442	21	AAAG10538
7	1498.5	61.5	442	21	AAAG32322
8	1498.5	61.5	442	21	AAAG40501
9	1498.5	61.5	473	21	AAAG10537

10	1498.5	61.5	473	21	AAAG40500
11	1498.5	61.5	474	21	AAAG32321
12	1493	61.3	466	21	AAAG5051
13	1480.5	60.8	437	16	AAAG2839
14	1439	59.1	469	16	AAAG2840
15	1425	58.5	460	23	ABAG3057
16	1402.5	57.6	400	21	AAAG10539
17	1402.5	57.6	400	21	AAAG32323
18	1402.5	57.6	400	21	AAAG40502
19	1385.5	56.9	397	23	ABAG3310
20	1349	55.4	471	16	AAAG2838
21	1349	55.4	471	16	AAAG6383
22	1157	47.5	407	22	ABAG69749
23	571	23.4	116	22	ABAG4416
24	571	23.4	116	22	ABAG7264
25	571	23.4	116	22	AAAG78158
26	571	23.4	116	22	AAAG22030
27	571	23.4	116	22	AAAG38402
28	571	23.4	116	23	ABAG47171
29	522	21.4	104	21	AAAG00253
30	362.5	14.9	430	22	AAAG83188
31	356	14.6	431	22	AAAG79515
32	356	14.6	437	22	AAAG93187
33	353	14.5	436	12	AAAG4357
34	353	14.5	437	22	AAAG83189
35	345.5	14.2	437	24	ABP79498
36	339.5	13.9	514	23	ABAG72726
37	338	13.9	509	23	ABAG2085
38	329	13.5	132	21	AAAG26766
39	325	13.3	101	21	AAAG26768
40	325	13.3	116	21	AAAG26767
41	320	13.1	134	22	AAAG00061
42	320	13.1	441	23	ABAG5981
43	315.5	13.0	427	21	AAAG15967
44	315.5	13.0	427	22	AAAG29342
45	310.5	12.7	489	22	AAAG67471

#### ALIGNMENTS

RESULT 1	
AAAG3432	AAAG3432 standard; Protein; 466 AA.
ID	
AAAG3432	
AC	AAAG3432;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human polypeptide, SEQ ID NO: 3064
XX	
KM	Human; full length cDNA; cDNA synthesis; oligo-capping.
XX	
OS	Homo sapiens.
XX	
PN	EP130094-A2.
XX	
PD	05-SEP-2001.
XX	
PF	07-JUL-2000; 2000EP-0114089.
XX	
PR	08-JUL-1999; 99JP-0194486.
XX	
PR	11-JAN-2000; 2000JP-0118774.
XX	
PR	02-MAY-2000; 2000JP-0183765.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX	
PI	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX	
DR	WPI: 2001-524255/58.
XX	
DR	N-PSDB: AAK94353.
XX	

Arabidopsis thalia  
Arabidopsis thalia  
Candida albicans p  
Sugar beet citrate  
Tobacco citrate SY  
S. cerevisiae BAX-  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
C. albicans BAX-as  
Potato citrate syn  
Potato citrate syn  
Drosophila melanog  
Peptide #11922 enc  
Protein #9263 enco  
Human bone marrow  
Peptide #8404 enco  
Peptide #12439 enc  
Human peptide enco  
Human secreted pro  
Corynebacterium th  
Corynebacterium gl  
C. glutamicum prote  
Citrate synthase e  
Brevibacterium lac  
N. gonorrhoeae aml  
Herbicidially activ  
Zea mays protein f  
Zea mays protein f  
Citrate synthase  
Lactococcus lactis  
E. coli proliferat  
Novel mar regulate  
Protonibacterium

PT 830 Primers useful for synthesizing full length cDNA clones and their  
 PT use in genetic manipulation -  
 XX  
 PS Claim 8; SEQ ID NO 3064; 1380bp + sequence listing; English.  
 CC  
 CC The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been  
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
 CC molecules have been determined. Primers for synthesizing the full length  
 CC cDNA are useful for clarifying the function of the protein encoded by  
 CC the cDNA. The full length clones were obtained by construction of full  
 CC length enriched cDNA libraries that were synthesised by the oligo-capping  
 CC method. The primers enable the production of the full length cDNA easily  
 CC without any special methods. The present sequence is a polypeptide  
 CC encoded by a full length human cDNA of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in CD-ROM format directly from EPO.  
 CC  
 XX  
 SQ Sequence 466 AA;  
 Query Match 100.0%; Score 2436; DB 22; Length 466;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-219;  
 Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALLTAARLGTGRKNSCLVLARHSAASSTNLKDLADLIPKQARIKTFROGHKTVV 60  
 DB 1 MALLTAARLGTGRKNSCLVLARHSAASSTNLKDLADLIPKQARIKTFROGHKTVV 60  
 QY 61 GQITVDMYMGKMGKGLVETSVLDDEGIRFGFSIPRCOKLLPRAKGEEPLPGLE 120  
 DB 61 GQITVDMYMGKMGKGLVETSVLDDEGIRFGFSIPRCOKLLPRAKGEEPLPGLE 120  
 QY 121 WLLVTGHIPTREOVSWLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAAYTALNSEN 180  
 DB 121 WLLVTGHIPTREOVSWLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAAYTALNSEN 180  
 QY 181 FARAAYAGISRTKYWEIYEDSMDLAKLPCVAAKIRNLYRBSGGIGALDSNLDGSHNF 240  
 DB 181 FARAAYAGISRTKYWEIYEDSMDLAKLPCVAAKIRNLYRBSGGIGALDSNLDGSHNF 240  
 QY 241 TNNLGYTDHOFTELTRLYTISHDHEGANSATSHLYGALSADPYISFAAMNGLAGPL 300  
 DB 241 TNNLGYTDHOFTELTRLYTISHDHEGANSATSHLYGALSADPYISFAAMNGLAGPL 300  
 QY 301 HGLANOEVLVMTLOLKEVKGKDVSDKLRDYIMNTLNSGSHVDPGYGHAVALRKTDPRYTCQ 360  
 DB 301 HGLANOEVLVMTLOLKEVKGKDVSDKLRDYIMNTLNSGSHVDPGYGHAVALRKTDPRYTCQ 360  
 QY 361 REEALKRLPMDPEKTLVAOLYKIVPNVLLLEOGKAKKNPVDVDAHSGVLLQYCGTETENNY 420  
 DB 361 REEALKRLPMDPEKTLVAOLYKIVPNVLLLEOGKAKKNPVDVDAHSGVLLQYCGTETENNY 420  
 QY 421 TVLFQVSRAVGVAQLQWLSRALGFLERPKSMSTEGIMRKTVDSKSG 466  
 DB 421 TVLFQVSRAVGVAQLQWLSRALGFLERPKSMSTEGIMRKTVDSKSG 466  
 RESULT 2  
 ABG32120  
 ID ABG32120 standard; Protein; 466 AA.  
 XX  
 AC ABG32120;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Human citrate synthase homologue (HUCS).  
 KW Human; citrate synthase homologue; HUCS; cancer; inflammatory disorder;  
 KW acquired immunodeficiency syndrome; AIDS; Addison's disease; allergy;  
 KW anaemia; asthma; atherosclerosis; bronchitis; Crohn's disease;  
 KW ulcerative colitis; atopic dermatitis; diabetes mellitus; emphysema;  
 KW glomerulonephritis; gout; Grave's disease; lupus erythematosus;  
 KW multiple sclerosis; myasthenia gravis; myocardial inflammation;

KW pericardial inflammation; osteoarthritis; osteoporosis; pancreatitis;  
 KW rheumatoid arthritis; scleroderma; Sjogren's syndrome; haemodialysis;  
 KW autoimmune thyroiditis; extracorporeal infection; circulation;  
 KW viral infection; bacterial infection; fungal infection; arrhythmia;  
 KW parasitic infection; protozoal infection; helminthic infection; trauma;  
 KW sympathetic nervous system disorder; angina; anaphylactic shock;  
 KW asthma; cardiovascular shock; Cushing's syndrome; hypertension;  
 KW hypoglycaemia; myocardial infarction; migraine; pheochromocytoma;  
 KW transgenic animal.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002099195-A1.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 27-NOV-2001; 2001US-0996223.  
 XX  
 PR 13-NOV-1997; 97US-0970134.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 PI Hillman JL, Lal P, Corley NC, Shah P;  
 DR WPI; 2002-666037/71.  
 DR N-PSDB; ABK90783.  
 XX  
 PS Claim 1; Fig 1A-F; 35pp; English.  
 CC  
 CC The invention describes an isolated human citrate synthase homologue  
 CC (HUCS) polypeptide (I). (I) is useful for screening a compound for  
 CC effectiveness as an agonist or antagonist useful for treating a disease  
 CC or condition associated with altered expression of functional HUCS in a  
 CC patient. (II), the encoding nucleic acid (II) and modulators of (I) are  
 CC useful for diagnosis, treatment and prevention of cancer, inflammatory  
 CC disorders such as acquired immunodeficiency syndrome (AIDS), Addison's  
 CC disease, allergies, anaemia, asthma, atherosclerosis, bronchitis, Crohn's  
 CC disease, ulcerative colitis, atopic dermatitis, diabetes mellitus,  
 CC emphysema, glomerulonephritis, gout, Grave's disease, lupus  
 CC erythematosus, multiple sclerosis, myasthenia gravis, myocardial or  
 CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,  
 CC rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune  
 CC thyroiditis, complications of cancer, haemodialysis and extracorporeal  
 CC circulation, viral, bacterial, fungal, parasitic, protozoal and  
 CC helminthic infections and trauma, and disorders of the sympathetic  
 CC nervous system including angina, anaphylactic shock, arrhythmias, asthma,  
 CC cardiovascular shock, Cushing's syndrome, hypertension, hypoglycaemia,  
 CC myocardial infarction, migraine and pheochromocytoma. (II) is useful for  
 CC creating knockin humanised animals or transgenic animals to model human  
 CC disease and to detect and quantify gene expression in biopsied tissues in  
 CC which expression of HUCS is correlated with disease. (II) is also useful  
 CC for generating hybridisation probes useful in mapping the naturally  
 CC occurring genomic sequence and oligonucleotide primers derived from  
 CC (II) are useful to detect single nucleotide polymorphisms. HUCS,  
 CC fragments of it and antibodies specific for HUCS are useful as elements  
 CC on a microarray which is useful to monitor or measure protein-protein  
 CC interactions, drug-target interactions and gene expression profiles. This  
 CC is the amino acid sequence of the human citrate synthase homologue  
 CC (HUCS).  
 CC  
 XX  
 SQ Sequence 466 AA;  
 Query Match 100.0%; Score 2436; DB 23; Length 466;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-219;  
 Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALLTAARLGTGRKNSCLVLARHSAASSTNLKDLADLIPKQARIKTFROGHKTVV 60  
 DB 1 MALLTAARLGTGRKNSCLVLARHSAASSTNLKDLADLIPKQARIKTFROGHKTVV 60

Qy	61	GOITVDMYTGMRGKMGKGLVETSVLDPBEGIRFREFSTIPBEOCKLLPRAKGGEETLPBGLF	120
Db	61	GOITVDMYTGMRGKMGKGLVETSVLDPBEGIRFREFGSIPEOCKLLPRAKGGEETLPBGLF	120
Qy	121	WLVVTGHIPTEBOVSWSLSEKMAKRAALBSHYVTMLDNPPTLJHMSOLSAAVPTALNSESN	180
Db	121	WLVVTGHIPTEBOVSWSLSEKMAKRAALBSHYVTMLDNPPTLJHMSOLSAAVPTALNSESN	180
Qy	181	FARAYAQGISRTKYWELIYEDSMDLIAKPCVAAKIYNNLIRESGIGALDINSIDMSHNF	240
Db	181	FARAYAQGISRTKYWELIYEDSMDLIAKPCVAAKIYNNLIRESGIGALDINSIDMSHNF	240
Qy	241	TNNLGTDTDHQFETLRLVLTITHSDEGCVNSAHTSHLVGSLSDPYISFAAAMNGLAGPL	300
Db	241	TNNLGTDTDHQFETLRLVLTITHSDEGCVNSAHTSHLVGSLSDPYISFAAAMNGLAGPL	300
Qy	301	HGLANOEVLWLTLOLQKEVGKDVSPSEKLRDYIMNTLNSGRVPCVGAHVLAKTDPRTYQ	360
Db	301	HGLANOEVLWLTLOLQKEVGKDVSPSEKLRDYIMNTLNSGRVPCVGAHVLAKTDPRTYQ	360
Qy	361	REFALKHLPNDPMPFKLYLAOLYKIVPNVLLDEQKAKPNPNDAHSGLVLLQYGTETENY	420
Db	361	REFALKHLPNDPMPFKLYLAOLYKIVPNVLLDEQKAKPNPNDAHSGLVLLQYGTETENY	420
Qy	421	TULFVYSRALVLAOLINSRALGFLPERPKSMTSEGLMKFVDSKSG	466
Db	421	TULFVYSRALVLAOLINSRALGFLPERPKSMTSEGLMKFVDSKSG	466

CC	XX	RESULT 3
CC	XX	ABB05627
CC	XX	ID ABB05627 standard; Protein; 466 AA.
CC	AC	ABB05627;
CC	DT	25-APR-2002 (first entry)
CC	XX	
CC	DE	Human citrate synthase homologue (HUCS) protein SEQ ID NO:1.
CC	XX	
CC	KW	Human; citrate synthase homologue; HUCS; enzyme; cytosolic; neotropic;
CC	XX	antiinflammatory; neuroprotective; citrate synthase; gene therapy;
CC	KW	cancer; inflammatory disorder; sympathetic nervous system disorder.
CC	XX	
CC	OS	Homo sapiens.
CC	XX	
CC	PN	US6323328-B1.
CC	PD	27-NOV-2001.
CC	XX	
CC	PE	13-NOV-1997; 97US-0970134.
CC	XX	
CC	PR	13-NOV-1997; 97US-0970134.
CC	XX	
CC	PA	(INCY-) INCYTE GENOMICS INC.
CC	XX	
CC	PI	Hillman JL, Lal P, Corley NC, Shah P;
CC	DR	WPI; 2002-163049/21.
CC	DR	N-PSDB; ABA93492.
CC	XX	
CC	PT	New polynucleotide encoding a human citrate synthase homolog for
CC	PT	preventing or treating cancer, inflammatory disorders and disorders of
CC	PT	the sympathetic nervous system
CC	PS	-
CC	XX	
CC	XX	Claim 1; Fig 1A-F; 31pp; English.
CC	XX	
CC	XX	This present sequence represents a human citrate synthase homologue
CC	XX	(HUCS) protein (I). (II), and the polynucleotide (II) encoding (I), have
CC	XX	cytosolic, antiinflammatory, neuroprotective and neotropic activities,
CC	XX	and can be used in gene therapy. (II) can be used to produce a HUCS
CC	XX	polypeptide. (I) and (II) can be used to prevent or treat disorders
CC	XX	associated with HUCS expression, particularly cancer, an inflammatory

CC.	disorder or a disorder of the sympathetic nervous system.
xx	
Sequence	466 AA:
Query Match	100.0%; Score 2436; DB 23; Length 466;
Best Local Similarity	100.0%; Pred. No. 2,7e-219;
Matches 466; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

OY	1	MALLTAAARLLGTRKNAKSCVLAARHSAASSTULXOJLADLIPKBEARIKFPROOHGTVV	60
Db	1	MALLTAAARLLGTRKNAKSCVLAARHSAASSTULXOJLADLIPKBEARIKFPROOHGTVV	60
OY	61	GOITVDMMYGGMGMGKGLVETSVLDPDEG1IFRRFOS1PECOKLLPKAKGGEEDLPBGLF	120
Db	61	GOITVDMMYGGMGMGKGLVETSVLDPDEG1IFRRFOS1PECOKLLPKAKGGEEDLPBGLF	120
OY	121	WLVTGHIPIPEEYVSWLSKEMAKRAALPSHYVTMLDNPPTNLHPKSOUSAAVTYLNEESN	180
Db	121	WLVTGHIPIPEEYVSWLSKEMAKRAALPSHYVTMLDNPPTNLHPKSOUSAAVTYLNEESN	180
OY	181	FARAYAOGISRTTYEMELIYEDSMIDLAKIPCAAKIYENLREGSGIGADISNIDMSHNF	240
Db	181	FARAYAOGISRTTYEMELIYEDSMIDLAKIPCAAKIYENLREGSGIGADISNIDMSHNF	240
OY	241	TNNLGYTDHOFTETRLRYLYTTHSDHEGCVNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL	300
Db	241	TNNLGYTDHOFTETRLRYLYTTHSDHEGCVNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL	300
OY	301	HGLANDEVIAWMLTQLOKEVKGXDVSDEKLBDYIWMNLNSGRVVPVGGHVAIVLKTTPRYTCQ	360
Db	301	HGLANDEVIAWMLTQLOKEVKGXDVSDEKLBDYIWMNLNSGRVVPVGGHVAIVLKTTPRYTCQ	360
OY	361	REFALKHLPNDPNPFKIVLAOLYKIVPVLVLEOGKAKNPNPNVDASHGVLLQYYGTENMY	420
Db	361	REFALKHLPNDPNPFKIVLAOLYKIVPVLVLEOGKAKNPNPNVDASHGVLLQYYGTENMY	420
OY	421	TUVPGVSRALGVLAOLINSRALGPTLERKSMSTGMLKPVDSKSG	466
Db	421	TUVPGVSRALGVLAOLINSRALGPTLERKSMSTGMLKPVDSKSG	466

XX	RESULT 4
ABG32121	
XX	ID ABG32121 standard; Protein; 464 AA.
XX	AC
XX	ABG32121;
DT	15-NOV-2002 (first entry)
XX	
DE	P1g citrate synthase homologue (HUCS).
XX	
KW	P1g; citrate synthase homologue; HUCS; cancer; inflammatory disorder;
KW	acquired immunodeficiency syndrome; AIDS; Addison's disease; allergy;
KW	anaemia; asthma; atherosclerosis; bronchitis; Crohn's disease;
KW	ulcerative colitis; atopic dermatitis; diabetes mellitus; emphysema;
KW	glomerulonephritis; gout; Grave's disease; lupus erythematosus;
KW	multiple sclerosis; myasthenia gravis; myocardial inflammation;
KW	pericardial inflammation; osteoarthritis; osteoporosis; pancreatitis;
KW	rheumatoid arthritis; scleroderma; Sjogren's syndrome; haemodialysis;
KW	autoimmune thyroiditis; extracorporeal infection; circulation;
KW	viral infection; bacterial infection; fungal infection; arrhythmia;
KW	parasitic infection; protozoal infection; helminthic infection; trauma
KW	sympathetic nervous system disorder; angina; anapylactic shock;
KW	asthma; cardiovascular shock; Cushing's syndrome; hypertension;
KW	hypoglycaemia; myocardial infarction; migraine; pheochromocytoma;
KW	transgenic animal.
XX	
OS	Sus scrofa.
XX	
PN	US200209195-A1.
XX	
PD	25-JUL-2002.

PF 27-NOV-2001; 2001US-0996223.  
 XX 13-NOV-1997; 97US-0970134.  
 XX (INCY-) INCYTE PHARM INC.  
 XX HILLMAN JL, Lal P, Corley NC, Shah P;  
 XX WPI; 2002-666037/71.  
 DR  
 XX  
 PT New human citrate synthase homolog polypeptide for diagnosing and  
 PT treating cancer, inflammatory disorders and disorders of the  
 PT sympathetic nervous system e.g. angina, and anaphylactic shock -  
 XX  
 PS  
 XX Disclosure; Fig 2A-B; 35pp; English.  
 CC The invention describes an isolated human citrate synthase homologue  
 CC (HUCS) polypeptide (I). (I) is useful for screening a compound for  
 CC effectiveness as an agonist or antagonist useful for treating a disease  
 CC or condition associated with altered expression of functional HUCS in a  
 CC patient. (II), the encoding nucleic acid (II) and modulators of (I) are  
 CC useful for diagnosis, treatment and prevention of cancer, inflammatory  
 CC disorders such as acquired immunodeficiency syndrome (AIDS), Addison's  
 CC disease, allergies, anaemia, asthma, atherosclerosis, bronchitis, Crohn's  
 CC disease, ulcerative colitis, atopic dermatitis, diabetes mellitus,  
 CC emphysema, glomerulonephritis, gout, Grave's disease, lupus  
 CC erythematosus, multiple sclerosis, myasthenia gravis, myocardial or  
 CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,  
 CC rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune  
 CC thyroiditis, complications of cancer, haemodialysis and extracorporeal  
 CC circulation, viral, bacterial, fungal, parasitic, protozoal and  
 CC helminthic infections and trauma, and disorders of the sympathetic  
 CC nervous system including angina, anaphylactic shock, arrhythmias, asthma,  
 CC cardiovascular shock, Cushing's syndrome, hypertension, hypoglycaemia,  
 CC myocardial infarction, migraine and pheochromocytoma. (II) is useful for  
 CC creating knockin humanised animals or transgenic animals to model human  
 CC disease and to detect and quantify gene expression in biopsied tissues in  
 CC which expression of HUCS is correlated with disease. (II) is also useful  
 CC for generating hybridisation probes useful in mapping the naturally  
 CC occurring genomic sequence and oligonucleotide primers derived from  
 CC (II) are useful to detect single nucleotide polymorphisms. HUCS,  
 CC fragments of it and antibodies specific for HUCS are useful as elements  
 CC on a microarray which is useful to monitor or measure protein-protein  
 CC interactions, drug-target interactions and gene expression profiles. This  
 CC is the amino acid sequence of the pig citrate synthase homologue (HUCS).  
 CC  
 XX  
 XX  
 SQ Sequence 464 AA;  
 Query Match 96.0%; Score 2339; DB 23; Length 464;  
 Best Local Similarity 96.1%; Pred. No. 3.3e-210;  
 Matches 446; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

DB 301 HGLANDEVVLWLTQLQKEVGKDVSDKLRDY INWTLNSGRVPGYGAVALRKTDPRYTCQ 360  
 QY 361 REFALKHLPNDPMEKTLVAQLKXKIVPNVLEOGKAKNPWVDASHGVLQYGMTEKNY 420  
 DB 361 REFALKHLPNDPMEKTLVAQLKXKIVPNVLEOGKAKNPWVDASHGVLQYGMTEKNY 420  
 QY 421 TVLFGVSRAVGVAQLIWSRALGPPLERPKSMSTEGIMKFDVSK 464  
 DB 421 TVLFGVSRAVGVAQLIWSRALGPPLERPKSMSTEGIMKFDVSK 464  
 RESULT 5  
 ABB60227  
 ID ABB60227 standard; Protein; 464 AA.  
 XX  
 AC ABB60227;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 7473.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL04330.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS  
 XX Disclosure; SEQ ID NO 7473; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 464 AA;  
 Query Match 71.3%; Score 1737.5; DB 22; Length 464;  
 Best Local Similarity 70.5%; Pred. No. 7.7e-154;  
 Matches 328; Conservative 50; Mismatches 80; Indels 7; Gaps 2;



Db 58 ARKMGETTIDMVGNGRIKALVETSVLDADEGINFRGLSIPCCOKVLPAADGTEPLP 117  
QY 117 ECLFMLLVYTHIPTEROVSMLSKEMAKRALPSHVYTMLDNPTNLHPMSQLSAAYTALN 176  
Db 118 EGLFMILLTGEVPTKSOVOOLSRMAERAPALPOHVYTMNNMFTLHPMSQFAAYTALN 177  
QY 177 SESNFARAYAOGISRTKRWELIYEDSMDLIAKLPCVAARIYRNLREGSGIGALDSDLMD 236  
Db 178 HDSKFAKAYSDGVHRSKSKWEYVEYEDSMDLIAKLPPVAATITTCYTGKSGSRIDSDLMD 237  
QY 237 SNHFTMLGTYDHOFEELTRLYITIHSDHGGVNSAHTSLYGALSIDPLSPAAAMNGL 296  
Db 238 SANFVKMLGYDAPFLMELRKLTYITHSDHGGVNSAHTSLYGALSIDPLSPAAALINGI 297  
QY 297 ACPRLHGLANQEVLYMVLTOLOKEVKGDSPEKLRDYMNTLNSGRVYPGYGHAYLRKTDPR 356  
Db 298 ACPRLHGLANQEVLYMVLRLQKEKGNPNSEQLKEYITWKILKSGQVYPGIGHAYLRKTDPR 357  
QY 357 YTCQREFALKHLPNDEMFRLVAQLKIVPNVLLDQKARNPMPNDAHSGVLLQYYGME 416  
Db 358 YTCQREFALKHLPEDETFLQVSKIKYKVPPILETGCKVKNPMPNDAHSGVLLQYYGME 417  
QY 417 MNYTYTLFGVSRLGVLQALINSRALGFLERPKSKSTEGIMKEV 461  
Db 418 MNYTYTLFGVSRLGVLASLWDRALGLPIERPKSFSTDLVKMV 462

RESULT 6

AAG10538  
ID AAG10538 standard; Protein: 442 AA.

AC AAG10538;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 8901.

KM Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PE 25-FEB-2000, 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145092.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.



PR 23-MAR-1999; 990S-0125788.  
PR 25-MAR-1999; 990S-0126264.  
PR 29-MAR-1999; 990S-0126785.  
PR 01-APR-1999; 990S-0127462.  
PR 06-APR-1999; 990S-0128234.  
PR 08-APR-1999; 990S-0128714.  
PR 16-APR-1999; 990S-0129845.  
PR 19-APR-1999; 990S-0130077.  
PR 21-APR-1999; 990S-0130449.  
PR 23-APR-1999; 990S-0130510.  
PR 23-APR-1999; 990S-0130891.  
PR 28-APR-1999; 990S-0131449.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 06-MAY-1999; 990S-0132487.  
PR 07-MAY-1999; 990S-0132863.  
PR 11-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134941.  
PR 19-MAY-1999; 990S-0134768.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 01-JUL-1999; 990S-0142154.  
PR 02-JUL-1999; 990S-0142055.  
PR 06-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142920.  
PR 12-JUL-1999; 990S-0142927.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 20-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 03-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.  
PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149368.  
PR 17-AUG-1999; 990S-0149175.  
PR 18-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PR 23-AUG-1999; 990S-0149902.  
PR 23-AUG-1999; 990S-0149930.  
PR 25-AUG-1999; 990S-0150566.  
PR 26-AUG-1999; 990S-0150884.  
PR 27-AUG-1999; 990S-0151065.  
PR 27-AUG-1999; 990S-0151066.  
PR 27-AUG-1999; 990S-0151080.  
PR 30-AUG-1999; 990S-0151303.  
PR 31-AUG-1999; 990S-0151438.  
PR 01-SEP-1999; 990S-0151430.  
PR 07-SEP-1999; 990S-0152363.  
PR 10-SEP-1999; 990S-0153070.  
PR 13-SEP-1999; 990S-0153758.  
PR 15-SEP-1999; 990S-0154018.  
PR 16-SEP-1999; 990S-0154039.  
PR 20-SEP-1999; 990S-0154779.  
PR 22-SEP-1999; 990S-0155139.  
PR 23-SEP-1999; 990S-0155486.  
PR 24-SEP-1999; 990S-0155659.  
PR 28-SEP-1999; 990S-0156458.  
PR 29-SEP-1999; 990S-0156596.  
PR 04-OCT-1999; 990S-0157117.  
PR 05-OCT-1999; 990S-0157753.  
PR 06-OCT-1999; 990S-0157865.

PR	07-OCT-1999:	99US-0158029.
PR	08-OCT-1999:	99US-0158232.
PR	12-OCT-1999:	99US-0158369.
PR	13-OCT-1999:	99US-0159293.
PR	13-OCT-1999:	99US-0159294.
PR	13-OCT-1999:	99US-0159295.
PR	14-OCT-1999:	99US-0159329.
PR	14-OCT-1999:	99US-0159330.
PR	14-OCT-1999:	99US-0159331.
PR	14-OCT-1999:	99US-0159637.
PR	18-OCT-1999:	99US-0159638.
PR	21-OCT-1999:	99US-0159584.
PR	21-OCT-1999:	99US-0160741.
PR	21-OCT-1999:	99US-0160767.
PR	21-OCT-1999:	99US-0160768.
PR	21-OCT-1999:	99US-0160770.
PR	21-OCT-1999:	99US-0160814.
PR	21-OCT-1999:	99US-0160815.
PR	22-OCT-1999:	99US-0160815.
PR	22-OCT-1999:	99US-0160980.
PR	22-OCT-1999:	99US-0160981.
PR	25-OCT-1999:	99US-0160989.
PR	25-OCT-1999:	99US-0161404.
PR	25-OCT-1999:	99US-0161405.
PR	25-OCT-1999:	99US-0161406.
PR	26-OCT-1999:	99US-0161359.
PR	26-OCT-1999:	99US-0161360.
PR	26-OCT-1999:	99US-0161361.
PR	28-OCT-1999:	99US-0161920.
PR	28-OCT-1999:	99US-0161920.
PR	28-OCT-1999:	99US-0161933.
PR	29-OCT-1999:	99US-0162142.

Query Match 61.5%; Score 1498.5; DB 21; Length 442;

Best Local Similarity 62.6%; Pred. No. 1,8e-131; Mismatches 83; Indels 1; Gaps 1;

QY	27	SASSTLKDILADLIREQOARIKTFROQKGTAVGQITVDMGKRGKGLVYETSVLD	86
DB	3	SSTDLKLSOLOLLIEQODRLKKLSEHGKVLGNTIDVMTGMRGKMTGLMTSTLSD	62
QY	87	PDEGIRFGFSIPECOKLLPKAKGGEPLPEGLFWLVTGHIPTREOVSWLSKEMAKRAA	146
DB	63	PERGIRFGFSIPECOKLLPTAOSGAPLEPEGLFWLVTGHIPTREOVSWLSKEMAKRAA	122
QY	147	LPESHVVTMDNFTNTHPSQLSAATVLTALNSENFPARAAYOGISRTKYWEITEDSMDLI	206
DB	123	VPOYVYNAIDALPSTAHPTQFASGVVALQVSEFQKAYENGHKSFEWETEDCLNLI	182
QY	207	AKLPCVAAKIRYRLYREGSGIGAIDSNLMDSHNFTNMLGVTDHQFELTFLYLTISHDHE	266
DB	183	ARPYVVAAYVRRMYKNGDSIPS-DKSLDYGANFSHMLGFDDDKVKEMLRLYITISHDHE	241
QY	267	GGNVAHSAHTSLVGSALSDPYLSFAAAMNGIAGPLHGLANOEVLWLTOLQKEVGKDVSD	326
DB	242	GGNVAHSAHTSLVGSALSDPYLSFAAAMNGIAGPLHGLANOEVLWLTOLQKEVGKDVSD	301
QY	327	KLRDYIWNLTNSGRVVPYGHANVLRKTDRTYTCQREFALKHLPNDPMEFKVADLYKIVPN	366
DB	302	QLKEVYWKTLNSGKVLPGYGHVLRNTDPRYVCOREFALKHLPNDPMEFKVADLYKIVPN	361
QY	387	VILLEGRKAKPMWNVNVAHSGVLLQYGMTEEMNYTTLFVGRALVLAOLIWSRALGFPL	446
DB	362	VLTLELCKVAKPMWNVNVAHSGVLLNHYGLTEARYITVLFVGRSLGICISOLIMDRALGLAL	421
QY	447	ERPKSMSTEGT 457	
DB	422	ERPKSVTMDWL 432	

RESULT 8  
AAG40501  
ID AAG40501 standard; Protein; 442 AA.  
XX

AC	AAG40501:	
XX		
DT	18-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 50261.	
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	termination sequence.	
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999:	99US-0121825.
PR	05-MAR-1999:	99US-0123180.
PR	09-MAR-1999:	99US-0123548.
PR	23-MAR-1999:	99US-0125788.
PR	25-MAR-1999:	99US-0126264.
PR	29-MAR-1999:	99US-0126785.
PR	01-APR-1999:	99US-0127462.
PR	06-APR-1999:	99US-0128234.
PR	08-APR-1999:	99US-0128714.
PR	16-APR-1999:	99US-0129845.
PR	19-APR-1999:	99US-0130077.
PR	21-APR-1999:	99US-0130449.
PR	23-APR-1999:	99US-0130510.
PR	23-APR-1999:	99US-0130891.
PR	28-APR-1999:	99US-0131449.
PR	30-APR-1999:	99US-0132048.
PR	30-APR-1999:	99US-0132407.
PR	04-MAY-1999:	99US-0132484.
PR	05-MAY-1999:	99US-0132485.
PR	06-MAY-1999:	99US-0132486.
PR	06-MAY-1999:	99US-0132487.
PR	07-MAY-1999:	99US-0132863.
PR	11-MAY-1999:	99US-0134256.
PR	14-MAY-1999:	99US-0134218.
PR	14-MAY-1999:	99US-0134219.
PR	14-MAY-1999:	99US-0134221.
PR	14-MAY-1999:	99US-0134370.
PR	18-MAY-1999:	99US-0134768.
PR	19-MAY-1999:	99US-0134941.
PR	20-MAY-1999:	99US-0135124.
PR	21-MAY-1999:	99US-0135353.
PR	21-MAY-1999:	99US-0135629.
PR	25-MAY-1999:	99US-0136021.
PR	27-MAY-1999:	99US-0136392.
PR	28-MAY-1999:	99US-0136782.
PR	01-JUN-1999:	99US-0137222.
PR	03-JUN-1999:	99US-0137528.
PR	04-JUN-1999:	99US-0137502.
PR	07-JUN-1999:	99US-0137724.
PR	08-JUN-1999:	99US-0138094.
PR	10-JUN-1999:	99US-0138540.
PR	10-JUN-1999:	99US-0138847.
PR	14-JUN-1999:	99US-0139119.
PR	16-JUN-1999:	99US-0139452.
PR	16-JUN-1999:	99US-0139453.
PR	17-JUN-1999:	99US-0139454.
PR	18-JUN-1999:	99US-0139455.
PR	18-JUN-1999:	99US-0139456.
PR	18-JUN-1999:	99US-0139457.
PR	18-JUN-1999:	99US-0139458.
PR	18-JUN-1999:	99US-0139459.
PR	18-JUN-1999:	99US-0139460.
PR	18-JUN-1999:	99US-0139461.
PR	18-JUN-1999:	99US-0139462.



QY 267 GGNVSAHTSLVGSALSDPYLSFAAMNGIAGPLHGLANOEVLWLTOLKEVGVKDYSD 326  
DB 242 GGNVSAHTGHLVGSALSDPYLSFAAMNGIAGPLHGLANOEVLIMIKSVVEEGEDISKE 301  
QY 327 KLRDYIWNLTNSGRVVGCHAVLRLKTDPRYTCOREFALHLPNDPMFKLVAOLYKIVPN 386  
DB 302 QLEKEYWKTINSKGVIPGIGHVLRNTDPRYVCOREFALHLPNDPMFKLVEVPP 361  
QY 387 VLEEGKAKNPMPNDVDAHSGVLLQYCGMTEMNYYTVLFVGRALGVLAQLIWSRACGFL 446  
DB 362 VLTELKVKMPMPNDVDAHSGVLLNHVGLTEARYTYTVLFVGRSLGICSQLIMBRALGLAL 421  
QY 447 ERKKSMTGCL 457  
DB 422 ERKKSMTMDWL 432

RESULT 9  
AAG10537  
ID AAG10537 standard; Protein; 473 AA.  
XX AAG10537;  
AC AAG10537;  
DT 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8900.  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144315.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147204.



PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145276.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145818.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 16-SEP-1999; 99US-0154779.  
PR 20-SEP-1999; 99US-0155139.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.





PR	01-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142055;
PR	06-JUL-1999;	9905-0142390;
PR	08-JUL-1999;	9905-0142820;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143524;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;
PR	16-JUL-1999;	9905-0144085;
PR	19-JUL-1999;	9905-0144086;
PR	19-JUL-1999;	9905-0144305;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144684;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145115;
PR	23-JUL-1999;	9905-0145248;
PR	23-JUL-1999;	9905-0145274;
PR	26-JUL-1999;	9905-0145226;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	28-JUL-1999;	9905-0145986;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147030;
PR	03-AUG-1999;	9905-0147034;
PR	04-AUG-1999;	9905-0147304;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148431;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149472;
PR	20-AUG-1999;	9905-0149723;
PR	23-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	25-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151083;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151530;
PR	07-SEP-1999;	9905-0151930;
PR	10-SEP-1999;	9905-0153707;
PR	13-SEP-1999;	9905-0153758;

PR	1-SEP-1999	9905-0154018
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0154479
PR	22-SEP-1999	9905-0155133
PR	23-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155659
PR	28-SEP-1999	9905-0156558
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157713
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	14-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159320
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	16-OCT-1999	9905-0159684
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160767
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160814
PR	21-OCT-1999	9905-0160815
PR	21-OCT-1999	9905-0160880
PR	22-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0160989
PR	23-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	25-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161359
PR	26-OCT-1999	9905-0161361
PR	28-OCT-1999	9905-0161320
PR	28-OCT-1999	9905-0161392
PR	28-OCT-1999	9905-0161943
PR	29-OCT-1999	9905-0162142

Query Match	61.58;	Score 1498.5;	DB 21;	Length 474;
Best Local Similarity	62.68;	Pred. No. 2e-131;		
Matches 270;	Conservative 77;	Mismatches 83;	Indels 1;	Gaps 1

QY	27	SASSNINLKDIADLADLPKROARKTKTPRQOHGKTVQVLIYDMYAGMRGKGLVETYSVD	86
Db	35	SSSTDIDLSQLOELLPEQODRLTKKSKSHGKQVLGNITVDWVIGMRGVTGLMTSTLSD	94
QY	87	PDEGIRFGFSPICECKLLPKAKGEEPLPELIFMLVYGHITPEEOVSMLSEWAKRAA	146
Db	95	PEGIRFREGSLTPEQOKVLPHTAQSAAEPLPEGLMLLLLTGKPKSPSEQVEALSKDLANRA	154
QY	147	LPSHVYVLMDFNPTNLPMSQLSAAVYTLNSESNNFARAYAOGISRTKXWELLYEDSMIL	206
Db	155	VPDYVYNNIDALPSTAHMPQTQFASCVMALOYSEQOKAYENGIRHKSFEWPEYEDCLNLT	214
QY	207	AKLPVAAKVIARNLTREGSGIGALDSNLDMSNFPNNMGYTPOHKTLETRLTITTHSHDE	266
Db	215	ARVPYVAAYVRRMKNKNDSTIPS-DKSIDYGNFNMHGLFDEDEKAKELMRLKLTITTHSHDE	273
QY	267	GGNSAHTSHLVGSALSDPYLSEFAAMNGLAGPLHGLANOEVLVLTOLQKEVKGDSDE	326
Db	274	GGNSAHTSHLVGSALSDPYLSEFAALNGLAGPLHGLANOEVLVLTITKSVSECGEDISKE	333
QY	327	KLRDVIWMTLNSGRVPGYGAHVLKCTPBRATYCOEFALKHLPNDMPKIVLAOKYIPN	386
Db	334	QJKEIYVMTLNSGKVIPIGTGIGVLTNTPDRYCOREFALKHLPDLPOLVLSKLEEVVP	393
QY	387	VLEEGKAKKNMPPNVDASHGVLTQYVGTENMYVYVLCVSRALCVLTAQLIWSRALGPPL	446









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2003, 20:10:52; Search time 61 Seconds  
(without alignments)  
1048.910 Million cell updates/sec

Title: US-09-996-223-1

Perfect score: 2436  
Sequence: 1 MALTYAARLLGTKNASCLV.....ERPMSSTGLKMFVDSKSG 466

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PTCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/PTCT\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US09C\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppa/US10C\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2436	100.0	466	US-09-996-223-1	Sequence 1, Appl
2	2239	96.0	464	US-09-996-223-3	Sequence 3, Appl
3	571	23.4	116	US-09-864-761-42562	Sequence 42562, A
4	356	14.6	437	US-09-738-626-6941	Sequence 6941, A
5	343.5	14.1	429	US-10-156-761-12864	Sequence 12864, A
6	325	13.3	1232	US-10-017-161-2396	Sequence 2396, Ap
7	320	13.1	134	US-09-734-569-80	Sequence 80, Appl
8	315.5	13.0	427	US-09-912-020-324	Sequence 324, Appl
9	274	11.2	426	US-09-815-242-11282	Sequence 11282, A
10	257.5	10.6	367	US-09-815-242-5427	Sequence 5427, Ap
11	257.5	10.6	373	US-09-815-242-12621	Sequence 12621, A
12	257.5	10.6	373	US-09-815-242-12906	Sequence 12906, A
13	256.5	10.5	383	US-09-738-626-4233	Sequence 4233, Ap
14	240.5	9.9	388	US-10-156-761-9964	Sequence 9964, Ap
15	237.5	9.7	394	US-09-738-626-4273	Sequence 4273, Ap

16	234.5	9.6	375	9	US-09-815-242-11803	Sequence 11803, A
17	227.5	9.3	389	9	US-09-815-242-10058	Sequence 10058, A
18	222.5	9.1	389	9	US-09-815-242-13894	Sequence 13894, A
19	183.5	7.5	366	15	US-10-156-761-11393	Sequence 11393, A
20	138	5.7	418	15	US-10-156-761-9965	Sequence 9965, Ap
21	125.5	5.2	608	10	US-09-344-882-12	Sequence 12, Appl
22	125.5	5.2	608	15	US-10-293-865-12	Sequence 12, Appl
23	114.5	4.7	272	10	US-09-344-882-10	Sequence 10, Appl
24	114.5	4.7	272	15	US-10-293-865-10	Sequence 10, Appl
25	100	4.1	363	10	US-09-925-300-1530	Sequence 1530, Ap
26	94.5	3.9	2092	15	US-10-128-714-3061	Sequence 3061, Ap
27	94.5	3.9	2092	15	US-10-128-714-8061	Sequence 8061, Ap
28	93	3.8	882	15	US-10-156-761-15031	Sequence 15031, A
29	93	3.8	904	15	US-10-156-761-8762	Sequence 8762, Ap
30	92.5	3.8	822	10	US-09-981-9008-9	Sequence 9, Appl
31	91	3.7	297	9	US-09-815-242-5206	Sequence 5206, Ap
32	91	3.7	601	15	US-10-156-761-12207	Sequence 12207, A
33	90.5	3.7	828	15	US-10-101-464A-934	Sequence 934, Ap
34	90.5	3.7	1039	12	US-10-021-660-80	Sequence 80, Appl
35	90	3.7	1168	10	US-09-919-603-2	Sequence 2, Appl
36	90	3.7	1172	9	US-09-919-770-4	Sequence 4, Appl
37	90	3.7	1172	12	US-10-301-822-203	Sequence 203, App
38	90	3.7	1172	15	US-10-060-036-171	Sequence 171, App
39	90	3.7	1172	15	US-10-020-141-14	Sequence 14, Appl
40	90	3.7	1172	15	US-10-017-724-2	Sequence 2, Appl
41	89.5	3.7	561	15	US-10-156-761-12820	Sequence 12820, A
42	88	3.6	1172	10	US-09-822-682-2	Sequence 2, Appl
43	88	3.6	4551	11	US-09-793-708-1	Sequence 1, Appl
44	88	3.6	4551	12	US-10-201-365-2	Sequence 2, Appl
45	88	3.6	4551	12	US-10-160-539-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-09-996-223-1

; Sequence 1, Application US/09996223  
; Patent No. US20020099195A1

GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.

LaI, Preeti  
Shah, Purni  
Corley, Neil C.

TITLE OF INVENTION: HUMAN CITRATE SYNTHASE HOMOLOG

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive  
CITY: Palo Alto

STATE: CA  
COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSPRO for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/996,223

FILING DATE: 27-NO. US20020099195A1-2001

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/970,134

FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0425 US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166  
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: ENDCNOT03  
CLONE: 2171653  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-996-223-1

Query Match 100.0%; Score 2436; DB 9; Length 466;  
Best Local Similarity 100.0%; Pred. No. 5.4e-237;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MALTTAAARLLGTAKNSCLVLAARHASASTNLKDLADLIRKEQARITFPOQHGKTV 60  
1 MALTTAAARLLGTAKNSCLVLAARHASASTNLKDLADLIRKEQARITFPOQHGKTV 60  
Qy 61 GOITVDMYGMKMGKGLVETSVLDPDEGIRFRGFSIPECOKLIPKAGGEPPLPEGLF 120  
61 GOITVDMYGMKMGKGLVETSVLDPDEGIRFRGFSIPECOKLIPKAGGEPPLPEGLF 120  
Db 121 WLLVGHITPEEQVSLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAAYTALNSESN 180  
121 WLLVGHITPEEQVSLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAAYTALNSESN 180  
Qy 181 PARAYAGISRTKRYMELIYEDSMDLAKLPCVAAKYRNLRYREGSGIGAIDSLDMSHNF 240  
181 PARAYAGISRTKRYMELIYEDSMDLAKLPCVAAKYRNLRYREGSGIGAIDSLDMSHNF 240  
Db 241 TNNLGYTDHQTFLRLYLTIHSDHEGNAVSAHTSHLVGSALSDPYLSFAAAMNGIAGPL 300  
241 TNNLGYTDHQTFLRLYLTIHSDHEGNAVSAHTSHLVGSALSDPYLSFAAAMNGIAGPL 300  
Qy 301 HGLANDEVVLWLTQLQKEVGKDVSDKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYTQ 360  
301 HGLANDEVVLWLTQLQKEVGKDVSDKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYTQ 360  
Db 361 REFALKHLNDMPFKVLAOLYKIVPNVLLQEGKAKPNPNVAHSGVLLQYGMTEMNYY 420  
361 REFALKHLNDMPFKVLAOLYKIVPNVLLQEGKAKPNPNVAHSGVLLQYGMTEMNYY 420  
Qy 421 TYLFGVSRAALGVLAOLINSRALGFLPERKSMSTEGIMKFDVSKG 466  
421 TYLFGVSRAALGVLAOLINSRALGFLPERKSMSTEGIMKFDVSKG 466  
Db 421 TYLFGVSRAALGVLAOLINSRALGFLPERKSMSTEGIMKFDVSKG 466

RESULT 2  
US-09-996-223-3  
Sequence 3, Application US/0996223  
Patent No. US20020099195A1  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Lal, Preeti  
Corley, Neil C.  
Shah, Puri  
TITLE OF INVENTION: HUMAN CITRATE SYNTHASE HOMOLOG  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/996,223  
FILING DATE: 27-Nov-02 US20020099195A1-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/970,134  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0425 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 464 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 164419  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-996-223-3

Query Match 96.0%; Score 2339; DB 9; Length 464;  
Best Local Similarity 96.1%; Pred. No. 3.3e-227;  
Matches 446; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MALTTAAARLLGTAKNSCLVLAARHASASTNLKDLADLIRKEQARITFPOQHGKTV 60  
1 MALTTAAARLLGTAKNSCLVLAARHASASTNLKDLADLIRKEQARITFPOQHGKTV 60  
Db 61 GOITVDMYGMKMGKGLVETSVLDPDEGIRFRGFSIPECOKLIPKAGGEPPLPEGLF 120  
61 GOITVDMYGMKMGKGLVETSVLDPDEGIRFRGFSIPECOKLIPKAGGEPPLPEGLF 120  
Db 121 WLLVGHITPEEQVSLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAAYTALNSESN 180  
121 WLLVGHITPEEQVSLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAAYTALNSESN 180  
Qy 181 PARAYAGISRTKRYMELIYEDSMDLAKLPCVAAKYRNLRYREGSGIGAIDSLDMSHNF 240  
181 PARAYAGISRTKRYMELIYEDSMDLAKLPCVAAKYRNLRYREGSGIGAIDSLDMSHNF 240  
Db 241 TNNLGYTDHQTFLRLYLTIHSDHEGNAVSAHTSHLVGSALSDPYLSFAAAMNGIAGPL 300  
241 TNNLGYTDHQTFLRLYLTIHSDHEGNAVSAHTSHLVGSALSDPYLSFAAAMNGIAGPL 300  
Qy 301 HGLANDEVVLWLTQLQKEVGKDVSDKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYTQ 360  
301 HGLANDEVVLWLTQLQKEVGKDVSDKLDYIMNTLNSGRVVPYGYHAVALRKTDPRYTQ 360  
Db 361 REFALKHLNDMPFKVLAOLYKIVPNVLLQEGKAKPNPNVAHSGVLLQYGMTEMNYY 420  
361 REFALKHLNDMPFKVLAOLYKIVPNVLLQEGKAKPNPNVAHSGVLLQYGMTEMNYY 420  
Qy 421 TYLFGVSRAALGVLAOLINSRALGFLPERKSMSTEGIMKFDVSK 464  
421 TYLFGVSRAALGVLAOLINSRALGFLPERKSMSTEGIMKFDVSK 464  
Db 421 TYLFGVSRAALGVLAOLINSRALGFLPERKSMSTEGIMKFDVSK 464

RESULT 3  
US-09-996-761-42562  
Sequence 42562, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO



```

: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
: SEQ ID NO 42562
: LENGTH: 116
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC020904.3
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EST_HUMAN HIT: BE901274.1, EVALUATE 1.00e-58
: OTHER INFORMATION: SWISSPROT HIT: O75390, EVALUATE 1.00e-61
US-09-864-761-42562

Query Match      23.4%: Score 571; DB 9; Length 116;
Best Local Similarity 92.2%: Pred. No. 1.2e-49;
Matches 107: Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      154  MLDNPTLHPMSOLAATVLTALNSFNARAYAGISRTKYWELIYEDSMDLIAKLPVCA 213
DB      1  MLDNPTLHPMSOLAATVLTALNSFNARAYAGISRTKYWELIYEDSMDLIAKLPVCA 60

QY      214  AKIYRNLYREGSGAIGALDSNLDWMSHNTNMLGYTDHQPTELRLVLTIHSDHEGN 269.
DB      61  AKIYRNLYREGSGAIGALDSNLDWMSHNTNMLGYTDHQPTELRLVLTIHSDHEGN 116

RESULT 4
US-09-738-626-6941
```

```

: Sequence 6941, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIRO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAOKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 6941
: LENGTH: 437
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-6941

Query Match      14.6%: Score 356; DB 10; Length 437;
Best Local Similarity 27.7%: Pred. No. 5e-27;
Matches 117: Conservative 83; Mismatches 149; Indels 74; Gaps 18;

QY      61  GQITVDMYGMGRCKGLVETSVLPDEGI-RFEGFSIPCCQKLLPRAKGEERLPEGL 119
DB      47  GLIFDPGVYSTGSTESKI---TYIDGAGILRYRGYDIAD---LAENATFNE-----V 94

QY      120  FWLVYTHIIPTEEOVSMLSKEMAKRAALPSHVYTMLDNPTNLHPMSOLAATVLTALNS 179
DB      95  SYLLINGELPPPDLEHKNDEIRHHTLLDEDFKSOFNVPDAPHPMATLASSVNLST-- 152

QY      180  NFARAYAGISRTKYWELIYEDSM-----LIAKLPVAAKIRNLYREGSG 226
DB      153  -----TYDQNLPLDEAQLDKATVLTAKVPMIAA--YAHARRKCAP 192

QY      227  IGALDSNLDWMSHNTNML-GYTDHQP-----TELRLVLTIHSDHEGNVSAHTSHLV 278
DB      193  YMPDPSNLNABENFLRMHFGVPTPEYELDPIMVKALDKL-LILHADHE-QNCSTSTVAMI 250

QY      279  GSALSDPLSLRAAMANGIAGPLAGLANOEVLVWLTOLQKEGKVSDEKLDIYIWNLTNS 338
DB      251  GSAOANMFVSJAGGINALSGPLHGANOVLEMLEDIKSNNGGATE--FMNKVKNKEDG 308

QY      339  GRVWPGYGHAVLRKTDPPYTCOREFA--LKHLPNDPFRKVAOLYKIVPVLLLEOGKAK 395
DB      309  VRLM-GFGHRYVKNYDPAALAYKETAHAILHLHSGDOLLDAIKLEIA--LADDFYFSR 365

QY      396  NPWPNVDAHSGVLLQYVGMTENNYTVLVGVSRALGVLA----OLIMSRALGFLEPRKS 451
DB      366  KLYPRVDVETGLIYRAMGF-PTDFETVYFAIGRLPGWIAHYREOL--GAAGKNKINRPQ 421

QY      452  MST 454
DB      422  VYT 424

RESULT 5
US-10-156-761-12864
: Sequence 12864, Application US/10156761
: Publication No. US20030119018A1
: GENERAL INFORMATION:
```

```

; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12864
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12864

```

```

Query Match          14.1%; Score 343.5; DB 15; Length 429;
Best Local Similarity 28.3%; Pred. No. 8,9e-26;
Matches 128; Conservative 67; Mismatches 140; Indels 117; Gaps 23;

```

```

QY 61 GQITVDMYGMGMKGLVYETSLDPEGI-RFRGFSIECOKLLPKAGGEPLE- 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 39 GLVTLDSGYNTAAYSKAI---TYLDEGQILRYGYP- 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 ---GLFWLVTGHIHPPEOVSMLSKEAKRAALPSHYVTLMDNPPNTHPMOLSAAVTA 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 82 TELEVAALLNGELPTVDELSTFKGETTQHTLHEDVKNFYRGPPRAHPMAMLSVSA 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 LNSESNFARAYAGISRTKYWEIYEDS-----MDLIAKPCVAAKIYRNLY 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 LST-----FYQDSNHPDEKQKNSTIRLAKLPITAAVAYK--- 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 222 RESSGICAIIDNSMDGSHFT---NMLGYTDHQTETL-----RLYLT 260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 ---KSIQ-----HPFYPRNDLGYVEN-FLRMFSVPAODYDLDPVVSALDKLI 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 261 IHSDEGNAVSAHTSHLVGSLDOPYLSFAAMNGLAGPLHGLANOEVLMLOLOKEVG 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 226 LHMADHE-QNCSTVRLVSSQANMFASISAGISALMGPLHGGNOSVLEMLBGIQTS-G 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 KDVSDEKLRDYINNTLNSGRVVPYGVHVLAKTDPRTTCOREFALKHL---PNDPMFKL 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 284 GDV-DPEIRKV--KKNKEDGVKLMGFGRHYVKNFDPRAKIIKAAAHVDLSLAKSDELDDI 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 377 VAQL--YKIYPNVLLDEGKAKKNPNDVAHSGVLQYGM-TEKNYTVLFGVSRLGVL 433
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 341 ALKLEHALADDFVE---RKLYPNVDFYTGILYRAMGPTPEW-FTVLFALGRLPGLWI 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 434 AOLIWSRAL---GPPLEPKSMST-EGIMKEV 461
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 395 AQ--WHEMIKEPSRIGRPROIYTGVEVLRDPV 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-10-017-161-2396
; Sequence 2396, Application US/10017161
; Publication No. US2003014366A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789

```

```

; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2396
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-2396

```

```

Query Match          13.3%; Score 325; DB 12; Length 1232;
Best Local Similarity 39.2%; Pred. No. 3.5e-23;
Matches 100; Conservative 16; Mismatches 79; Indels 60; Gaps 8;

```

```

QY 104 LEPKAGGEBPLP-----EGLEFWLVTGHIPT---EQO---VSWSLSEMAKRAALPSH 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 467 MLPALLRPEPLPTPTPVHLLQGLGLLGGGLMLAITLLEERLLPVWSLSEMAKRAALPSH 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 151 VVTMLDNFPNTHPMOLSAAVTAALSESNFARAYAGISRTKYWEIYEDSMDLIAK 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 527 VVTMLDNFPNTHPMOLSAAVTAALSESNFARAYAGISRTKYWEIYEDSMDLIAK 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 208 ---KLPCVAAKIYRNLYREGSGAIDNSLDMSHNFTNMLGYTDHQTETLRLYTIH 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 587 TLSFSRLSTLSTRRRGLMCSGSGCSLSLP-----LPPHQIPEIARHGFTYC 637
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 263 SDHEGNAVSAHTSH---LVG-----SALSDPLISFAAMNGIA 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 638 GRPALNOCGOREGARRPVGLRRPOLARRGLPPTPVKTLRALKGPALGPALREAT 697
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 GPLHGLANOEVLMV 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 698 GPCALP-----FWL 707
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
US-09-734-569-80
; Sequence 80, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lersch, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reinold, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv
; FILE REFERENCE: BASE-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patentln Ver. 2.1/Mordperfect 6.1
; SEQ ID NO 80
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-80

```

```

Query Match          13.1%; Score 320; DB 9; Length 134;
Best Local Similarity 61.9%; Pred. No. 3.2e-24;
Matches 60; Conservative 16; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 28 ASSTNLKDLADLIPREOARIKTRROOHGKTVVGOITVDVAMGGMKGLVYETSLDP 87
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 38 SKTVDMKERRALAEILPEODLKKIKKDYKISLGDITVDMCIGMGISGMLMETSLLDA 97
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Oy 88 DEGRFSGFSIPECOKLLPKAKGEEPLFGLFWLLV 124  
Db 98 DEGRFSGFSIPECOKLLPKAISGEEPLFGLFWLLV 134

RESULT 8  
US-09-912-020-324

Sequence 324, Application US/09912020  
Patent No. US20020045592A1

GENERAL INFORMATION:

APPLICANT: Zyskind, Judith  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Twissick, John  
APPLICANT: Forsyth, R. Allyn  
APPLICANT: Froelich, Jamie M.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
FILE REFERENCE: ELITRA.001DVI  
CURRENT APPLICATION NUMBER: US/09/912,020  
CURRENT FILING DATE: 2001-07-23  
PRIOR APPLICATION NUMBER: 09/492,709  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/117,405  
PRIOR FILING DATE: 1999-01-27  
NUMBER OF SEQ ID NOS: 485  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 324  
LENGTH: 427  
TYPE: PRT  
ORGANISM: E. COLI  
US-09-912-020-324

Query Match 13.0%, Score 315.5; DB 9; Length 427;  
Best Local Similarity 26.4%; Pred. No. 5.9e-23;  
Matches 116; Conservative 76; Mismatches 160; Indels 87; Gaps 20;

Oy 57 KTVVGQITVDMYMGMRM-----KGLVETG-----VLDDDEGRF-RGFSIPECOKLL 105  
Db 22 KGTLDQVDYDIRTSGKVFETFDPEFTSTASCESKIRPFIDDEGLLHGRFPI--DGLA 78  
Oy 106 PKAKGEEPLFGLFWLLVTHGHPTEEQVSMLEKAKRAALPSHVVTMLNFPNTLHPM 165  
Db 79 TDSNVLE-----VCYLLNGEKPTQEOYDEKFTVTHTMTIHQITLPLFAFRDSHPM 132  
Oy 166 SOLSAAVTALNSESFARAYAOISRTKYMELIYEDSM-----LIAKLPCV 212  
Db 133 AVMGITGALAA-----FYHDSLDVNNPRHREIAFRLLSKMPMT 172  
Oy 213 AAKTYRNLYREGSGIGALDSNLDMSHNTNMLGTDHOFTELRL-----YLTHSDH 265  
Db 173 AAMCYK--YSTGQPPVYPRNDLSYAGNFMNFSTPCPYEVNPLERAMDRIILLHADH 230  
Oy 266 EGVNVSATSHLVGSALSDPYLSFAAANGLAFLHGLANOEVWLWLTOL--KEVGADV 323  
Db 231 E-QNASTSTVTAASSGANPPACIAAGIASLMGPRAGHGANEAALMLEIISVYKHIFEV 289  
Oy 324 SDEKLRDYIWNLTNSGRVYPGYGAHVLKRTDPRYTCOREF--ALKHL-PNDMPKFLVAQ 379  
Db 290 RRADKN-----DSFLRM-GFGHRYVKNYDPRATVMRETCHEVYKELGTADDLENAME 342  
Oy 380 LYKTYPN--VLLEGGKAKNPWPNDANHSGLVLLQYGMTEMYTYLVLEVSALGVLAOLI 437  
Db 343 LENTLALNDPYFIE--KRLYPNDVDFSGIILKAMGPS-SMFTVIFAMARTGVIAH-- 395  
Oy 438 WS--FALGFLPERPKSNST 454  
Db 396 WSEMHSDGMKTIARPROLYT 414

RESULT 9

US-09-815-242-11282  
Sequence 11282, Application US/09815242  
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Twissick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11282  
LENGTH: 426  
TYPE: PRT  
ORGANISM: Helicobacter pylori  
US-09-815-242-11282

Query Match 11.2%, Score 274; DB 9; Length 426;  
Best Local Similarity 27.7%; Pred. No. 9.1e-19;  
Matches 108; Conservative 72; Mismatches 144; Indels 66; Gaps 18;

Oy 72 MRGKGLVETGSDVLDPEGRFSGFSIPECOKLLPKAKGEEPLFGLFWLLVTHGHPTE 131  
Db 59 VNGKKELY-----YRGHRT--EDLVAKTKYVD-----VCRLLTGELPKN 97  
Oy 132 EOVSWM-LSKEMAKRAALPSHVVTMLDNEPTNLMHMSOLSAAVTALNSESFARAYAOIS 190  
Db 98 ODESLEFELRLHRSFVHESLNMFSAPPSNAHPMAKLSGVSILST-----LYSTH 149  
Oy 191 RTKYMELIYED-SMDLIAKLPCVAAKIYRNLYREGSGIGALDSNLDMSHNTNML-GY-- 246  
Db 150 QNMHTVEEYQTMARIYAKIPTLAICYRN--EYGAPIIYDPIARSYENTLFLMRGPY 207  
Oy 247 -----TDHOFTELRL-----YLTHSDHEGVNVSATSHLVGSALSDPYLSFAAAMN 294  
Db 208 SRLKHTTQGEVEIITPLEVEARDKITLTLHADH--SQNASSTYRNVASTGVHYYAIAISGIS 266  
Oy 295 GLAFLHGLANOEVWLWLTOL--KEVGADVSDEKLRDYIWNLTNSGRVYPGYGAHVLK 352  
Db 267 ALMGHLHGAGANEKVLQLEBIGDVKNVDKYIA--RVKD-----KNDNFKLMGFGHRYVKS 319  
Oy 353 TDPRYTCOREFALK--HLPDMPKFLVAOLYKYPNVVL--EOGKAKNPWPNDANHSGLV 407  
Db 320 YDPRAKTIK--GLDDELHQKGVKMKDERSETIAAVEETALKDEYFIERNLVYPNVDYSGT 377  
Oy 408 LLOYGKTEMYTYLVLEVSALGVLAOLI 437  
Db 378 ILRALKI-PVRFETPVYIGRTVGMCAOLL 406

RESULT 10

```
US-09-815-242-5427
; Sequence 3427, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5427
; LENGTH: 367
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5427
```

```
Query Match 10.6%; Score 257.5; DB 9; Length 367;
Best Local Similarity 26.6%; Pred. No. 3.3e-17;
Matches 100; Conservative 71; Mismatches 156; Indels 49; Gaps 15;

QY 73 RGMKGLV-YETSVLDPEG-IRFRGFSIPBCQKLPKAKGGEPLDEGLFWLVGHIPT 130
DB 1 RGLGVIYAETKISSIIESQITVAGYDIDL-----AENAOFEVIFLLMNYRLPN 51
QY 131 EEOVSLSKEMAKRALPSHVYTMIDNFT-NLHPMSQSAVTALNSESNFARAAGI 189
DB 52 EEBLAHLKGLNDYMTLNPRVYTHFEERYVDHVPMTALRTSLSYI-----AHFPD 103
QY 190 SRTKWEELIYEDSMELIAKPCVAARIYRLYREGSGIGAIDSNLDMSHNFTMLG---Y 246
DB 104 AENESDENRERARMIOAKVASLVTAFAF--VRQDKEPLKPNPDLSYAANFLMLGELP 161
QY 247 TDHOFELIRLYTTHSDHEGNAVSAHTSHLVGSALSDPYLSFAAAMNGIAGPLHGLANO 306
DB 162 TDIE-VEAFNKALILHADHE-LNASFARCAVSSISDMYSGIYAAGVSLKGLHGGANE 219
QY 307 EVLVWLTQLOKEVGKDVSDKLDYIMNTLNSGRVYPGYGHAVLRKTDPRYTCOREFA-- 364
DB 220 QVNTMIS---ETG---STENVADYLDKFRANKDKVGFGRHYKKGDDPRAKYLREMSRO 272
QY 365 -LKLHPNDPMFKLVAQLYKIVPVNLEOGKAKNPMPNVAHSGVLLQYGMTEM--NYVT 421
DB 273 ITKDAGREELFEKSVKMEK-----RMAEEKGLINVDYFSA---SVYHCMELPHDLFT 322
QY 422 VLEGVSRALGVLAQLI 437
DB 323 PIRAVSRASGMIHAIL 338
```

RESULT 11

```
US-09-815-242-12621
; Sequence 12621, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12621
; LENGTH: 373
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-12621
```

```
Query Match 10.6%; Score 257.5; DB 9; Length 373;
Best Local Similarity 26.6%; Pred. No. 3.4e-17;
Matches 100; Conservative 71; Mismatches 156; Indels 49; Gaps 15;

QY 73 RGMKGLV-YETSVLDPEG-IRFRGFSIPBCQKLPKAKGGEPLDEGLFWLVGHIPT 130
DB 6 RGLGVIYAETKISSIIESQITVAGYDIDL-----AENAOFEVIFLLMNYRLPN 56
QY 131 EEOVSLSKEMAKRALPSHVYTMIDNFT-NLHPMSQSAVTALNSESNFARAAGI 189
DB 57 EEBLAHLKGLNDYMTLNPRVYTHFEERYVDHVPMTALRTSLSYI-----AHFPD 108
QY 190 SRTKWEELIYEDSMELIAKPCVAARIYRLYREGSGIGAIDSNLDMSHNFTMLG---Y 246
DB 109 AENESDENRERARMIOAKVASLVTAFAF--VRQDKEPLKPNPDLSYAANFLMLGELP 166
QY 247 TDHOFELIRLYTTHSDHEGNAVSAHTSHLVGSALSDPYLSFAAAMNGIAGPLHGLANO 306
DB 167 TDIE-VEAFNKALILHADHE-LNASFARCAVSSISDMYSGIYAAGVSLKGLHGGANE 224
QY 307 EVLVWLTQLOKEVGKDVSDKLDYIMNTLNSGRVYPGYGHAVLRKTDPRYTCOREFA-- 364
DB 225 QVNTMIS---ETG---STENVADYLDKFRANKDKVGFGRHYKKGDDPRAKYLREMSRO 277
QY 365 -LKLHPNDPMFKLVAQLYKIVPVNLEOGKAKNPMPNVAHSGVLLQYGMTEM--NYVT 421
DB 278 ITKDAGREELFEKSVKMEK-----RMAEEKGLINVDYFSA---SVYHCMELPHDLFT 327
QY 422 VLEGVSRALGVLAQLI 437
DB 328 PIRAVSRASGMIHAIL 343
```

RESULT 12

US-09-815-242-12906

Sequence 12906, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12906

LENGTH: 373

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-12906

Query Match 10.6%; Score 257.5; DB 9; Length 373;

Best Local Similarity 26.6%; Pred. No. 3,4e-17;

Matches 100; Conservative 71; Mismatches 156; Indels 49; Gaps 15;

QY 73 RGMKGLVY-ETSVLDPDDEG-IRFRGFSIPECOKLPPKAKGGEPLPEGLFWLVTGHIPT 130

DB 6 RGLGVIAAEFKISSIISQLYAGYDIDL-----AENAFEEVIFLLMYRLPN 56

QY 131 EBOVSWLSKEMAKRAALPSHVYTMLDNFP--NLHPMSQLSAAVTALNSENFRARYAOGI 189

DB 57 EEEHLNHLGKLNQYMTLNPVYTHFEYVTDHVMFTALRTSLSYI-----AHFDPD 108

QY 190 SRTKMWELIYEDSMOILAKLCVAAKIYRNLYREGSGIGALIDSNLDMSHNFTNMG---Y 246

DB 109 AENEDENRYERFAMKIOAKVASLVTAFAR--VRQKEPLKPNPDLSYANFELYMLRGELP 166

QY 247 TDHOFTELRLYLTIHSDHEGNNVSAHTSHLVGSLSDPYLSFAAAMGLAGPLHLGLANQ 306

DB 167 TDIE-VEAFNKALLIHADHE-LMASAFARCAVSSLSDMYSIGVIAAAGSLGRLPGANE 224

QY 307 EYLVWLTQLOKEVKGQVDEKLDYIMNTLNSGRVYVPGYGAVALKKTDPRTYCOREFA-- 364

DB 225 QVMYMLS---EIG---SIENVDAYLDEKFAKDKVMGFGHRYVYKDGDPRAKYLRMSRO 277

QY 365 -LKHLPNDPMFKLVNOLKRYIPNVLLBOGKAKNPNPANDAHSGVLLQYGMTEM--NYT 421

DB 278 ITKDAKREBELPEMSYMEK-----RMAEEKGLIPNDVYSY---SVYHCHMEIPHLFT 327

QY 422 VLEGVSRALGVLAQLI 437

DB 328 PIFAVSRSGAGWTATHL 343

RESULT 13

US-09-738-626-4233

Sequence 4233, Application US/09738626

Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAMA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIRO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NAOKO

APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO

APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738,626

CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 4233

LENGTH: 383

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

US-09-738-626-4233

Query Match 10.5%; Score 256.5; DB 10; Length 383;

Best Local Similarity 24.3%; Pred. No. 4.5e-17;

Matches 94; Conservative 81; Mismatches 147; Indels 65; Gaps 14;

QY 73 RGMKGLVY-ETSVLDPDDEGIRFRGFSIPECOKLPPKAKGGEPLPEGLFWLVTGHI 128

DB 11 KGLGVIAADYAVASVMEFNTSLYRGVAVEDLV-----ENCSPFEVIFLLMHGEL 61

QY 129 PTEBOVSWLSKEMAKRAALPSHVYTMLDNFPNLHPMSQLSAAVTALNSENFRARYAOG 188

DB 62 PTAQOLAEFNEGRGYSRLDAGLISLHSLPEKAMPDMVMTAVSYM-----G 109

QY 189 ISRTKYW---ELIYEDSMOILAKLCVAAKIYRNLYREGSGIGALIDSNLDMSHNFTNML 244

DB 110 TKDSEYFTTDESHIKKVGHTLLAQLPMLADIR--RKGLDIAPDSKSYAENLMSV 167

QY 245 GYT-----DHOFTELRLYLTIHSDHEGNNVSAHTSHLVGSLSDPYLSFAAAMN 294

DB 168 FGTGESPASNPADVDREKES---LILVAEH-SFNASFTFARVITSTKSDVYSATGTAIG 223

QY 295 GIAGPLHGLANDEVYLVLTQLOKEVKGQVDEKLDYIMNTLNSGRVYVPGYGAVALKKTDP 354

DB 224 ALKGPLHGLGANEFVHTMLAID-----DPNKAAMVNNMLDNKNVYMGFGHRYKKRGD 276

QY 355 PRYTQO-----REFALKHLPNDPMFKLVNOLKRYIPNVLLBOGKAKNPNPANDAHSGVLLQ 410

DB 277 SRVPSMEKSFRELAARHGE---KWVA-MIENMDADARIGIK---PNLDFPAGRAYH 328

QY 411 YVGMTEMYVTVLEGVSRALGVLAQLI 437

DB 329 LLGF-PVDFFPLFVIARVAGWTATIV 354

RESULT 14

US-10-156-761-9964

Sequence 9964, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIOYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156, 761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 9964  
LENGTH: 388  
TYPE: PRN  
ORGANISM: Streptomyces avermitilis  
US-10-156-761-9964

Query Match 9.98; Score 240.5; DB 15; Length 388;  
Best Local Similarly 24.18; Pred. No. 1.9e-15;  
Matches 98; Conservative 67; Mismatches 172; Indels 69; Gaps 16;

QY 59 VVGQITVDMYGGKRGKGLVETSVLDPEGIRFGFSIPEQOKLLPKAKGGEPLPEG 118  
DB 17 LAGVVTDTSLGDVKGLEGFYH-----YROYSAVE---LAOTRGEED----- 55  
QY 119 LFWLVTGHIPTREQ-----VSLSKEMAKRAALPSSHVTMLDNPPTNLHPMSOL 168  
DB 56 VVHLMVHVGELPDQRAVFTARTALRLPD--GVRLTLAVAVAA-----AGRGPLAGL 108  
QY 169 SAAVTALNSESFPRAAQAQISRTKYWELIYEDSMDLAKLPCVAAKTYRNLYREGSGIG 228  
DB 109 RTALSLGAAKGFPPVY--DIDR---EORRADVMAAFAVPTLLTA---LVRLGQGLD 158  
QY 229 AID--SNLDWSHNTNML--GYTDHQTTELRLVLTIHSDHEGGSVAHSHLVGSALSD 284  
DB 159 PVERRELSAANTLYMLTSEPDARTAEQYLSTIDH--GNASTFTARTVTSIGAD 217  
QY 285 PYSFAAMNGLAGPLHGLANOEVWLTLQKEVGKDVSDKLRDYIMNTLNSGRVVP 344  
DB 218 VAAALVAGVAGLSGFLHGAPSRALDTLDAIG-----TPERDSIMIREVLAGDRIMG 270  
QY 345 YGHAVALKTPRTTCQREPAKLPNDPMKLVLAQLYKYPNVLLGEGAKNPVNDAAH 404  
DB 271 FGHVYVTEDEPRSMRLGVAEQF--GGPLVDFAVEVRHVEALIAELKPGRELHTNVEFY 328  
QY 405 SGVLLQYGMTENMYTVLFGVSRALGVLAQLIMSRALGPPLEPK 450  
DB 329 AGVVMELGILPR-EMFTPTFAARVVG-----WSANILQAEDEPK 367

RESULT 15  
US-09-738-626-4273  
Sequence 4273, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIALI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, MOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738, 626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 4273  
LENGTH: 394  
TYPE: PRN  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4273

Query Match 9.78; Score 237.5; DB 10; Length 394;  
Best Local Similarly 23.18; Pred. No. 3.9e-15;  
Matches 90; Conservative 74; Mismatches 155; Indels 71; Gaps 16;

QY 73 RGKKGGLYE---TSVLDPEGIRFGFSIPEQOKLLPKAKGGEPLPEGFLVLTGHI 128  
DB 21 KGLNGVISTDTSISKWAPESNSLTYKQYAVEDLY-----ENCSEFEVITLLWFGEL 71  
QY 129 PTEEOVSLSKEMAKRAALPSSHVTMLDNPPTNLHPMSOLSAVTAALS-----ESN 180  
DB 72 PTEEQLRTEFTTGRSVRSLDAGLSLSHSLPNTCHPMVDLRTAVSYMGTFDPPTRDAD 131  
QY 181 PARAYAOGISRTKYWELIYEDSMDLAKLPCVAAKTYRNLYREGSGIGALDSNLDWSHNF 240  
DB 132 HIRSIQH-----NLALQLEPMVAVAMDIR--RSGEELIAPDHKKGIASNF 173  
QY 241 TNNLGYTD-----HQTFLTRILVLTIHSDHEGGSVAHSHLVGSALSDPYLSFAA 292  
DB 174 LSNVFGNDGSVANSADIDFER--SLILYAEH--SFNASTFSARVLSSTRSDIYSAITGA 231  
QY 293 MNGIAGPLHGLANOEVWLTLQKEVGKDVSD--EKLRYIMNTLNSGRVVPYGHAVLR 351  
DB 232 IGALGPLHGGANE--FVMTML-----DIDDPNNAADMCKALDRKERIMGFGHRYK 283  
QY 352 KTRPRTY---TCQREPAKLPNDPMKLVLAQLYKYPNVLLGEGAKNPVNDAAHSGV 407  
DB 284 NGDSRVPSMEKSMKSLAARHGO---KWV--HMYESMOEVEARTGIR---PMLDPAPG 335  
QY 408 LQYVGMTEKNYTVLFGVSRALGVLAQLI 437  
DB 336 AYYMLGF-PVDEFTPLFLVLAIVSGWTAHIY 364

Search completed: September 5, 2003, 20:20:24  
Job time: 64 secs

Check  
JP

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2003, 20:06:12 : Search time 29 seconds  
(without alignments)  
679.892 Million cell updates/sec

Title: US-09-996-223-1

Perfect score: 2436

Sequence: 1 MALTTAARLGTKNASCLV.....ERPKSMTEGLMKFYDSKSG 466

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfilltest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2436	100.0	466	US-08-970-134-1	Sequence 1, Appl
2	2339	96.0	464	US-08-970-134-3	Sequence 3, Appl
3	338	13.9	431	US-09-252-991A-25969	Sequence 25969, A
4	314	12.9	424	US-09-328-352-4187	Sequence 4187, Ap
5	259	10.6	386	US-09-134-001C-3770	Sequence 3770, Ap
6	234.5	9.6	390	US-09-252-991A-29273	Sequence 29273, A
7	217	8.9	391	US-09-328-352-4861	Sequence 4861, Ap
8	100	4.1	1118	US-09-585-173B-36	Sequence 36, Appl
9	96	3.9	1447	US-09-376-330-17	Sequence 17, Appl
10	94.5	3.9	3665	US-08-222-617A-13	Sequence 13, Appl
11	94.5	3.9	3712	US-08-222-617A-4	Sequence 4, Appl
12	94.5	3.9	3712	US-08-222-617A-25	Sequence 25, Appl
13	92	3.8	1705	US-08-669-765-4	Sequence 4, Appl
14	91	3.7	353	US-09-252-991A-22881	Sequence 22881, A
15	90	3.7	1172	US-08-313-288B-19	Sequence 19, Appl
16	89	3.7	1489	US-08-669-785-2	Sequence 2, Appl
17	89	3.7	1706	US-08-669-785-2	Sequence 2, Appl
18	89	3.7	1794	US-09-328-352-5522	Sequence 32058, A
19	88	3.6	269	US-09-252-991A-32058	Sequence 2, Appl
20	88	3.6	459	US-09-252-991A-32058	Sequence 2, Appl
21	88	3.6	912	US-08-951-871-2	Sequence 1, Appl
22	88	3.6	4551	US-09-320-878-1	Sequence 1, Appl
23	88	3.6	4551	US-09-141-908-2	Sequence 1, Appl
24	88	3.6	4551	US-09-657-440-1	Sequence 31, Appl
25	88	3.6	4613	US-09-105-537-31	Sequence 6, Appl
26	88	3.6	11877	US-09-105-537-6	Sequence 7, Appl
27	87.5	3.6	461	US-09-894-698-7	Sequence 2, Appl

28	87.5	3.6	483	US-09-894-698-2	Sequence 2, Appl
29	87.5	3.6	888	US-08-953-492-2	Sequence 2, Appl
30	87.5	3.6	988	US-08-286-819A-19	Sequence 19, Appl
31	87.5	3.6	988	US-08-980-357-19	Sequence 19, Appl
32	87.5	3.6	999	US-08-473-553A-5	Sequence 5, Appl
33	87	3.6	1031	US-09-585-173B-40	Sequence 40, Appl
34	87	3.6	1168	US-08-448-170-6	Sequence 6, Appl
35	87	3.6	1168	US-08-961-803-8	Sequence 8, Appl
36	86.5	3.6	700	US-07-862-588B-2	Sequence 2, Appl
37	86	3.5	635	US-08-873-479-43	Sequence 43, Appl
38	86	3.5	659	US-08-136-277-3	Sequence 3, Appl
39	86	3.5	659	US-08-479-403-3	Sequence 3, Appl
40	86	3.5	659	US-08-835-734-3	Sequence 3, Appl
41	86	3.5	679	US-09-252-991A-29548	Sequence 29548, A
42	85.5	3.5	552	US-09-801-191A-6	Sequence 6, Appl
43	85.5	3.5	702	US-09-328-352-8176	Sequence 8176, Ap
44	85.5	3.5	828	US-09-107-532A-4191	Sequence 4191, Ap
45	85.5	3.5	1183	US-08-447-031A-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-970-134-1  
Sequence 1, Application US/08970134  
Patent No. 6323328  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: HUMAN CITRATE SYNTHASE HOMOLOG  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/970,134  
FILING DATE: Herewith  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0425 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: ENDCMOT03  
CLONE: 2171653  
US-08-970-134-1  
Query Match 100.0%; Score 2436; DB 4; Length 466;

Best Local Similarity 100.0%; Pred. No. 2.8e-242; Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALLTAAARLLGTRKNSCLVLAARHSAASSTNLKDLADLIPREQARITFRQOHGTYV 60  
 DB 1 MALLTAAARLLGTRKNSCLVLAARHSAASSTNLKDLADLIPREQARITFRQOHGTYV 60  
 QY 61 GQITVDMYGGMRGMKGLVYETSVLDPDDEGIRFGFSIPECOKLLPKAGGEEPLEGTLF 120  
 DB 61 GQITVDMYGGMRGMKGLVYETSVLDPDDEGIRFGFSIPECOKLLPKAGGEEPLEGTLF 120  
 QY 121 WLLVTGHIPTREOVSWLSKEMAKRALPSHVVTMLDNFPTNLHPMSQLSAATALNSESN 180  
 DB 121 WLLVTGHIPTREOVSWLSKEMAKRALPSHVVTMLDNFPTNLHPMSQLSAATALNSESN 180  
 QY 181 FARAYAGISRTKYWEIYEDSMDLAKLPCVAAKIYRNLYREGSGIGALDSNLDMSHNF 240  
 DB 181 FARAYAGISRTKYWEIYEDSMDLAKLPCVAAKIYRNLYREGSGIGALDSNLDMSHNF 240  
 QY 241 TNNLGYTDHOFTELRLYLTTHSDHEGNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300  
 DB 241 TNNLGYTDHOFTELRLYLTTHSDHEGNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300  
 QY 301 HGLANOEVLWMLQLOKEVGKDVSDKLRDYIMNTLNSGRVPGYGAVALRKTDPRTYTCQ 360  
 DB 301 HGLANOEVLWMLQLOKEVGKDVSDKLRDYIMNTLNSGRVPGYGAVALRKTDPRTYTCQ 360  
 QY 361 REFALKHLPNDPMEKLVKIQYKIPVNLLEOGKAKNPWPVDAHSGVLLQYGMTEMNYY 420  
 DB 361 REFALKHLPNDPMEKLVKIQYKIPVNLLEOGKAKNPWPVDAHSGVLLQYGMTEMNYY 420  
 QY 421 TVLFGVSRAIGVLAQLIWSRALGFLPERKSMSTEGIMKRFVDSKSG 466  
 DB 421 TVLFGVSRAIGVLAQLIWSRALGFLPERKSMSTEGIMKRFVDSKSG 466

## RESULT 2

US-08-970-134-3  
 : Sequence 3, Application US/08970134  
 : Patent No. 6323328

GENERAL INFORMATION:  
 : APPLICANT: Hillman, Jennifer L.  
 : APPLICANT: Lal, Preeti  
 : APPLICANT: Corley, Neil C.  
 : TITLE OF INVENTION: HUMAN CITRATE SYNTHASE HOMOLOG  
 : NUMBER OF SEQUENCES: 3  
 : CORRESPONDENCE ADDRESSES:  
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 : STREET: 3174 Porter Drive  
 : CITY: Palo Alto  
 : STATE: CA  
 : COUNTRY: USA  
 : ZIP: 94304

COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette  
 : OPERATING SYSTEM: DOS  
 : SOFTWARE: FastSeq for Windows Version 2.0  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/970,134  
 : FILING DATE: Herewith  
 : CLASSIFICATION: 530  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER:  
 : FILING DATE:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Billings, Lucy J.  
 : REGISTRATION NUMBER: 36,749  
 : REFERENCE/DOCKET NUMBER: PF-0425 US  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 650-855-0555  
 : TELEFAX: 650-845-4166

TELEX:  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 464 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : IMMEDIATE SOURCE:  
 : LIBRARY: GenBank  
 : CLONE: 164419  
 : US-08-970-134-3

Query Match 96.0%; Score 2339; DB 4; Length 464;  
 Best Local Similarity 96.1%; Pred. No. 2.8e-232; Matches 446; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 MALLTAAARLLGTRKNSCLVLAARHSAASSTNLKDLADLIPREQARITFRQOHGTYV 60  
 DB 1 MALLTAAARLLGTRKNSCLVLAARHSAASSTNLKDLADLIPREQARITFRQOHGTYV 60  
 QY 61 GQITVDMYGGMRGMKGLVYETSVLDPDDEGIRFGFSIPECOKLLPKAGGEEPLEGTLF 120  
 DB 61 GQITVDMYGGMRGMKGLVYETSVLDPDDEGIRFGFSIPECOKLLPKAGGEEPLEGTLF 120  
 QY 121 WLLVTGHIPTREOVSWLSKEMAKRALPSHVVTMLDNFPTNLHPMSQLSAATALNSESN 180  
 DB 121 WLLVTGHIPTREOVSWLSKEMAKRALPSHVVTMLDNFPTNLHPMSQLSAATALNSESN 180  
 QY 181 FARAYAGISRTKYWEIYEDSMDLAKLPCVAAKIYRNLYREGSGIGALDSNLDMSHNF 240  
 DB 181 FARAYAGISRTKYWEIYEDSMDLAKLPCVAAKIYRNLYREGSGIGALDSNLDMSHNF 240  
 QY 241 TNNLGYTDHOFTELRLYLTTHSDHEGNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300  
 DB 241 TNNLGYTDHOFTELRLYLTTHSDHEGNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300  
 QY 301 HGLANOEVLWMLQLOKEVGKDVSDKLRDYIMNTLNSGRVPGYGAVALRKTDPRTYTCQ 360  
 DB 301 HGLANOEVLWMLQLOKEVGKDVSDKLRDYIMNTLNSGRVPGYGAVALRKTDPRTYTCQ 360  
 QY 361 REFALKHLPNDPMEKLVKIQYKIPVNLLEOGKAKNPWPVDAHSGVLLQYGMTEMNYY 420  
 DB 361 REFALKHLPNDPMEKLVKIQYKIPVNLLEOGKAKNPWPVDAHSGVLLQYGMTEMNYY 420  
 QY 421 TVLFGVSRAIGVLAQLIWSRALGFLPERKSMSTEGIMKRFVDSK 464  
 DB 421 TVLFGVSRAIGVLAQLIWSRALGFLPERKSMSTEGIMKRFVDSK 464

## RESULT 3

US-09-252-991A-25969  
 : Sequence 25969, Application US/09252991A  
 : Patent No. 6551795

GENERAL INFORMATION:  
 : APPLICANT: Marc J. Rubenfield et al.  
 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 : FILE REFERENCE: 107196.136  
 : CURRENT APPLICATION NUMBER: US/09/252,991A  
 : PRIOR FILING DATE: 1999-02-18  
 : PRIOR APPLICATION NUMBER: US 60/074,788  
 : PRIOR FILING DATE: 1998-02-18  
 : PRIOR APPLICATION NUMBER: US 60/094,190  
 : PRIOR FILING DATE: 1998-07-27  
 : NUMBER OF SEQ ID NOS: 33142  
 : SEQ ID NO 25969

LENGTH: 431  
 : TYPE: PRT  
 : ORGANISM: Pseudomonas aeruginosa  
 : US-09-252-991A-25969

Query Match 13.9%; Score 338; DB 4; Length 431;  
 Best Local Similarity 27.0%; Pred. No. 4.4e-26;



[illegible]

	Query Match	10.6%	Score 259;	DB 4;	length 386;
	Best Local Similarity	26.3%	Pred. No. 5,1e-18;		
	Matches 102;	Conservative 69;	Mismatches 177;	Indels 40;	Gaps 14
OY	55	HGKTVVGQITVDMMTGMRGMKGLVYETSVLD-PDEGIRERFGSIPCCOKLLPRKAGGEE	113		
		:  :			
Db	4	HKEVFGKFKIMELKRGLEGV--IAAEFKISSIIDISQLTGAYGIDDL-----AEN	52		
OY	114	PLPEGFMLLYVGHITPEBOVSMLSKEMAKGAALPSHYVYTMDFPT-NLHPMSOLSAAY	172		
		:  :			
Db	53	AOPEELIIFLLMNYRLCPNESELSLKKLPDYMTLNRRYKHFEEYVDVHNHPITALTSTV	112		
OY	173	TALNSESFNARAYAAGISRTKYMWELIVEDSMDLIAKPCVAAKIYRNILYREGSGAIDS	232		
		:  :			
Db	113	SYV-----AHPPPEAKNESDEKKYDRAIRIQAKIASLVYTNAR--VADGKEPPKPS	162		
OY	233	NLDMSINPTNNML-GYTDHOF-TELTRLVYLTHSDHEGNAVSAHTSHLVGSALSDPLSTRA	290		
		:  :			
Db	163	ELSYAANFLYLMRGELPTEVEVEAFENKALLIHAHDE-IMASAFPARCAVSLSDMDSGVV	221		
OY	291	AAANGLAGPLHGLANOEVLYMVLTOLOKRVGVDGDEKLRDYINWYTLMSGVAVVGYCHAVL	350		
		:  :			
Db	222	AALGSLKGLPHGGADEKRWMSL----KELG---SIDVNDHITLDERFANKKIKINGFGRYY	274		
OY	351	RKTDPYTCQREFALKHLPPNDPMFKLVYQLKVIYPNVLLDEGKAKKPNPVDYAHSGVLQ	410		
		:  :			
Db	275	KDDPRAKYLREMSKRITETEGQSELFEMLSLAIEKRMRKEEGKL---PVDVFSATV--	328		
OY	411	YEGMT-EMYYTYVLFVGRALGVLAQLI	457		
		:  :			
Db	329	YHSMNIPHDLPPIFAVSRTSGWIAHL	356		



Db 588 TVGRNIEDVYR-GSKGLIAD-----VELL-----ROITEASGAIITFEVERTNNKG 634  
OY 270 ---VSANTSHLVGSALSDPYLSFAAAMNGLAGPLHGLANOEVLWLTQLOKEVGKDVSD 325  
Db 635 QVVDVTKNLSTMLGGLSEPMVQHLSTKKEFIRADEKL--KTLFPLFGECLKLI---VAD 689  
OY 326 EKLRDYIMNTLNGSNVPGYHVAVLKRTDPRYTCOREFAL--KHLPNDEMFELVAQLYKI 383  
Db 690 NELGS-LKIALEGSYVEPGDPIRNPKVLPTGKNIHALDPOAIPTTAAKLSA-----KI 744  
OY 384 VPNVLLEOGKAKN--PMPVNDVHSGVLLQYGTMTENMTYTVFGVSRLGVLAQLIIM 438  
Db 745 IVDRLEROKVDNGKYPET-----IALVLMGTDNIKTY-----GES-----LAQYLM 787

RESULT 9  
US-09-376-330-17  
Sequence 17, Application US/09376330  
Patent No. 6399321  
GENERAL INFORMATION:  
APPLICANT: Tessier, Daniel C.  
APPLICANT: Dignard, Daniel  
APPLICANT: Bergeron, John J.M.  
TITLE OF INVENTION: Method for screening for  
TITLE OF INVENTION: UDP-glucose:glycoprotein glucosyltransferase (UGGT) activity  
FILE REFERENCE: 2139-9"us"  
CURRENT APPLICATION NUMBER: US/09/376,330  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 17  
LENGTH: 1447  
TYPE: PRT  
FEATURE:  
ARTIFACT: Artificial Sequence  
OTHER INFORMATION: S. pombe UGCT  
US-09-376-330-17

Query Match 3.9%; Score 96; DB 4; Length 1447;  
Best Local Similarity 20.8%; Pred. No. 2.9;  
Matches 103; Conservative 62; Mismatches 177; Indels 154; Gaps 23;

OY 10 LIGTKNAS-CLVLAARHASASTNKLADLPKEQARITFRQGHKTAVGQATTVM 68  
Db 675 LIGSSNAKYSFWLVADEFNEKEGLELISLADLLSENKDNAMLMLOEGNNHVPPLFAKL 734  
OY 69 YGMRGMRKGL-VYETSVLPDEGIRFGRFSIPECOKLPRKAKGEEPLFGLF-----L 122  
Db 735 SSPKSSKHLEILNSSLDPSSGV---VNDMDKALKFLKSKAVVKE--GLTGCKSAL 789  
OY 123 LVYGH-----IPTEQVSWLSK---EMAKRAALPSHVVTMLDNFPTN 161  
Db 790 LINGMICSFSVDSLNTADLMKLMQMEYDNLKSNAGSSRLKNSRAISFLSSYLKT 849  
OY 162 LH--PMQOLS-----AAVTALNSESNEPARAYNOGIS-----RTKYW 195  
Db 850 LESTPMTSTSPPTKEEKLFPDPFIYNNKLCVGNATFETDPSKAYYQFVAVLPLDSKDSOKW 909  
OY 196 ELIYEDSMDLIAPCVNAKTYRNLVREGSGIGALDSIMDSHNFTNMLGYTDHOFTEL 255  
Db 910 SAL-----LEAVSKLNGVGRVRIHLNPKQTLSEL-----PLT 940  
OY 256 RLYLTIHSDHGGCVNASTSHLVGSALSDPYLSFAAAMNGLAGPLHGLANOEVL---VWL 312  
Db 941 RFRYR-----SISAEPEFDALCHLEESYVEF---DNL--PADTLITMDIEARDAM- 985  
OY 313 TOLQKEVGKDVSEKLT-----RDYIMNTLNGSNVPGYHVAVLKRTDPR----- 356  
Db 986 TVMKQDVVIDLFPNKLKLEHTSEALDSTFAIYELKNT--LVQGSQEEFRKSPRGMOLOK 1043  
OY 357 -----YTCOREFALKHLPN-----DPMFKLVAQLYKI VPNVLLEOGK 393

Db 1044 LGNLTNSHYTDITVLNGLY-----FOLKANGVLTLEPMGRSSQFYEILS--LNKNS 1096  
OY 394 AKNPMPNVDHSGVLL 409  
Db 1097 YKDPQVIVDSEGVTL 1112

RESULT 10  
US-08-222-617A-13  
Sequence 13, Application US/08222617A  
Patent No. 5882879  
GENERAL INFORMATION:  
APPLICANT: Veenstra, Annemarie E.  
APPLICANT: Martin, Juan F.  
APPLICANT: Garcia, Bruno D.  
APPLICANT: Gutierrez, Santiago  
APPLICANT: Barredo, Jose L.  
APPLICANT: Von Doehren, Hans  
APPLICANT: Palissa, Harriet  
APPLICANT: Van Lierp, Henk  
TITLE OF INVENTION: A Method for Influencing Beta-Lactam  
TITLE OF INVENTION: Antibiotic Production and for Isolation of Large  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/222,617A  
APPLICATION NUMBER: US/08/222,617A  
FILING DATE: 04-Apr-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
REFERENCE/DOCKET NUMBER: 97,157  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3665 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Acremonium chrysogenum  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..3665  
OTHER INFORMATION: /label= ACVS  
OTHER INFORMATION: /note= "ACV synthetase from Acremonium  
OTHER INFORMATION: chrysogenum; aa 1-3665"  
US-08-222-617A-13

Query Match 3.9%; Score 94.5; DB 2; Length 3665;  
Best Local Similarity 21.7%; Pred. No. 19;  
Matches 63; Conservative 47; Mismatches 91; Indels 89; Gaps 15;

OY 60 VGOITVDMVGMGMKGLVYETSVLPDEGIRFGRFSIPECOKLPRKAKGEEPLFGL 119  
Db 671 LGRADFOIKLNGIIEPEIESTLAMYP--GIR---ASTVYSKLLSQ--GQETIOHL 722  
OY 120 --FWILVYGHIPTEQVSWLSKEWAKRAALPSHVVTMLDNFPTNHL-----PMQOLS 169  
Db 120 --FWILVYGHIPTEQVSWLSKEWAKRAALPSHVVTMLDNFPTNHL-----PMQOLS 169

Db 723 VGIYVDEGHIEBGLSLFLEKK-LPRYVWPTRLV-QLAQIPNTNGKADLRALPAVEYA 780  
QY 170 AAVT-----ALNSESFARAYAGISRTKRWELIYED 201  
Db 781 VAPTHKODGERGNQLESIDLAIWGNILSVPAODIGSESNEFRUGSHIA-----829  
QY 202 SMDLIAPKPCVAAKIYRNLYREGSGIGALDSNIDMSHNTNMLGYTDHOFTELRLYLTI 261  
Db 830 CIOIARV-----ROOLGOGITLLEEVFOTKLRAMAALLSEKRYTKAS-----871  
QY 262 HSDHEGCGN-VSAHTSHLVGSA-----LSDPYLSFAAMNGLAGPLHGLANO 306  
Db 872 -----NQTNGVTNGTAHVNGHANGHVSYSYA-SSLQCGFV--YHSLKNE 914

RESULT 11  
US-08-222-617A-4  
; Sequence 4, Application US/08222617A  
; Patent No. 5882879  
; GENERAL INFORMATION:  
; APPLICANT: Veenstra, Annemarie E.  
; APPLICANT: Martin, Juan F.  
; APPLICANT: Garcia, Bruno D.  
; APPLICANT: Gutierrez, Santiago  
; APPLICANT: Barredo, Jose L.  
; APPLICANT: Von Doehren, Hans  
; APPLICANT: Palissa, Harriet  
; APPLICANT: Montenegro, Eduardo P.  
; TITLE OF INVENTION: A method for influencing Beta-lactam  
; TITLE OF INVENTION: Antibiotic Production and for isolation of large  
; TITLE OF INVENTION: Quantities of ACV Synthetase  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,617A  
; FILING DATE: 04-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; REFERENCE/DOCKET NUMBER: 97,157  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; NAME/KEY: Modified-site  
; LOCATION: 2555  
; OTHER INFORMATION:  
; OTHER INFORMATION: /note="Xaa-Ala or Ser"  
US-08-222-617A-4

Query Match 3.9%; Score 94.5; DB 2; Length 3712;  
Best Local Similarity 21.7%; Pred. No. 20;  
Matches 63; Conservative 47; Mismatches 91; Indels 89; Gaps 15;  
QY 60 VGOITVDMYGMGRKGLVETSVDPDEGIRFRGFSIPECOKLLPKAKGGEPLPEGL 119  
Db 671 LGRADQIKIRIGRIEPIESTLAMP--GIR---ASIVSKKLSQ---GQETIQDHL 722  
QY 120 --FWLLVTGHIPIEEOVSWLSKEMAKRALPSHVYTMIDNFPNINH-----PMSQLS 169

Db 723 VGIYVDEGHIEBGLSLFLEKK-LPRYVWPTRLV-QLAQIPNTNGKADLRALPAVEYA 780  
QY 170 AAVT-----ALNSESFARAYAGISRTKRWELIYED 201  
Db 781 VAPTHKODGERGNQLESIDLAIWGNILSVPAODIGSESNEFRUGSHIA-----829  
QY 202 SMDLIAPKPCVAAKIYRNLYREGSGIGALDSNIDMSHNTNMLGYTDHOFTELRLYLTI 261  
Db 830 CIOIARV-----ROOLGOGITLLEEVFOTKLRAMAALLSEKRYTKAS-----871  
QY 262 HSDHEGCGN-VSAHTSHLVGSA-----LSDPYLSFAAMNGLAGPLHGLANO 306  
Db 872 -----NQTNGVTNGTAHVNGHANGHVSYSYA-SSLQCGFV--YHSLKNE 914

RESULT 12  
US-08-222-617A-25  
; Sequence 25, Application US/08222617A  
; Patent No. 5882879  
; GENERAL INFORMATION:  
; APPLICANT: Veenstra, Annemarie E.  
; APPLICANT: Martin, Juan F.  
; APPLICANT: Garcia, Bruno D.  
; APPLICANT: Gutierrez, Santiago  
; APPLICANT: Barredo, Jose L.  
; APPLICANT: Von Doehren, Hans  
; APPLICANT: Palissa, Harriet  
; APPLICANT: Montenegro, Eduardo P.  
; TITLE OF INVENTION: A method for influencing Beta-lactam  
; TITLE OF INVENTION: Antibiotic Production and for isolation of large  
; TITLE OF INVENTION: Quantities of ACV Synthetase  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,617A  
; FILING DATE: 04-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; REFERENCE/DOCKET NUMBER: 97,157  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-222-617A-25

Query Match 3.9%; Score 94.5; DB 2; Length 3712;  
Best Local Similarity 21.7%; Pred. No. 20;  
Matches 63; Conservative 47; Mismatches 91; Indels 89; Gaps 15;  
QY 60 VGOITVDMYGMGRKGLVETSVDPDEGIRFRGFSIPECOKLLPKAKGGEPLPEGL 119  
Db 671 LGRADQIKIRIGRIEPIESTLAMP--GIR---ASIVSKKLSQ---GQETIQDHL 722  
QY 120 --FWLLVTGHIPIEEOVSWLSKEMAKRALPSHVYTMIDNFPNINH-----PMSQLS 169  
Db 723 VGIYVDEGHIEBGLSLFLEKK-LPRYVWPTRLV-QLAQIPNTNGKADLRALPAVEYA 780  
QY 170 AAVT-----ALNSESFARAYAGISRTKRWELIYED 201



ADDRESSSEE: Cooper & Dunham LLP  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version 4.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/313,288B  
 FILING DATE: January 5, 1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 278-0400  
 TELEFAX: (212) 391-0526  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1172 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-313-288B-19

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2003, 20:06:07 : Search time 101 Seconds  
(without alignments)  
1190.619 Million cell updates/sec

Title: US-09-996-223-1  
Perfect score: 2436  
Sequence: 1 MALLTAARILGTKNASCLV.....ERPKSMSTEGIMKFDKSG 466

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPRMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	100.0	466	4	Q96FZ8
2	2323	95.4	466	11	Q8VHF5
3	2321	95.3	466	11	Q9CZU6
4	2219	91.1	466	11	Q9DAM4
5	1737.5	71.3	464	5	Q9W401
6	1735	71.2	522	5	Q8T477
7	1549	63.6	474	3	Q8NKF2
8	1539.5	63.2	474	3	Q8P8D6
9	1507	61.9	472	3	Q9HFT1
10	1501.5	61.6	469	3	Q9P688
11	1497.5	61.5	456	3	Q8N170
12	1483.5	60.9	460	5	Q8SSP7
13	1481.5	60.8	472	10	Q9FUT7
14	1480.5	60.8	437	3	Q96544
15	1480.5	60.8	475	3	Q9P4E3
16	1473.5	60.5	433	10	Q9M1D3

17	1457.5	59.8	473	10	Q946X8	Q946X8 prunus pers
18	1439	59.1	469	10	Q24135	Q24135 nicotiana t
19	1396	57.3	264	4	Q9BWN8	Q9BWN8 homo sapien
20	1377.5	56.5	493	5	Q9XCN7	Q9XCN7 oryza sativ
21	1232	50.6	478	5	Q95T24	Q95T24 oryza sativ
22	1158	47.5	460	8	Q9TEM3	Q9TEM3 emericella
23	1108.5	45.5	339	10	Q24259	Q24259 populus bal
24	1071	44.0	261	5	Q8T9J9	Q8T9J9 drosophila
25	931.5	38.2	563	5	Q816W7	Q816W7 plasmidium
26	685	28.1	138	6	Q8WMY1	Q8WMY1 bos taurus
27	580	23.8	188	10	Q8H9E3	Q8H9E3 sesdania ro
28	366.5	15.0	470	16	Q8RQ3	Q8RQ3 corynabacte
29	346	14.2	429	16	Q9R339	Q9R339 streptomyce
30	345.5	14.2	427	16	Q9J0X0	Q9J0X0 neisseria m
31	341	14.0	373	17	Q974S5	Q974S5 sulfolobus
32	339.5	13.9	514	10	Q9LX56	Q9LX56 arabidopsis
33	338	13.9	509	10	Q9S0H7	Q9S0H7 arabidopsis
34	336.5	13.8	377	2	Q9LX9	Q9LX9 thermus the
35	336.5	13.8	412	2	Q9KUF8	Q9KUF8 rickettsia
36	335.5	13.8	429	16	Q8F801	Q8F801 leprospira
37	333	13.7	413	2	Q9F111	Q9F111 rickettsia
38	332.5	13.6	377	2	Q8VL23	Q8VL23 thermus aqu
39	331.5	13.6	410	2	Q9LAF3	Q9LAF3 rickettsia
40	331.5	13.6	410	2	Q9LAF5	Q9LAF5 rickettsia
41	331.5	13.6	411	2	Q8KW7	Q8KW7 rickettsia
42	331.5	13.6	431	16	Q33066	Q33066 mycobacteri
43	331	13.6	413	2	Q9F110	Q9F110 rickettsia
44	330.5	13.6	411	2	Q59459	Q59459 israeli tic
45	330	13.5	411	2	Q8KW8	Q8KW8 rickettsia

## ALIGNMENTS

### RESULT 1

Q96FZ8 ID Q96FZ8 PRELIMINARY: PRT: 466 AA.  
AC Q96FZ8:  
DT 01-DEC-2001 (TRMBLrel. 19, Created)  
DT 01-DEC-2001 (TRMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TRMBLrel. 23, Last annotation update)  
DE Hypothetical protein FLJ90475 (EC 4.1.3.7) (Citrate synthase).  
DE Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T., Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K., Masuko Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N., Hattori A., Okumura K., Iwayanagi T., Niinomiya K.;  
RT \*NBD human cDNA sequencing project.\*;  
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O + OXALOACETATE.  
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
CC EMBL: BC010106; AAH10106.1; -;  
CC EMBL: AK074956; BAC11314.1; -;  
CC InterPro: IPR002020; Citrate\_synth.  
CC Pfam: PF00285; citrate\_synth.  
CC PRINTS: PR00143; CITRATESYNTHASE.  
DR PROSITE: PS00480; CITRATE\_SYNTHASE; 1.  
DR Hypothetical protein; Lyase; Tricarboxylic acid cycle.  
KW SEQUENCE 466 AA; 51712 MW; 459CB29C0BA06997 CRC64;  
SQ

```

Query Match      100.0%; Score 2436; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.8e-190;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALLTAARLLGKNNKSCVLARHNASASTNKLIDLADLPKEQARITFRQOHGKTYV 60
DB 1 MALLTAARLLGKNNKSCVLARHNASASTNKLIDLADLPKEQARITFRQOHGKTYV 60
QY 61 GOITVDMYGMKGMKGLYETSVLDPDGIIFRGFSIPECOKLIPKAKGGEPLPEGIF 120
DB 61 GOITVDMYGMKGMKGLYETSVLDPDGIIFRGFSIPECOKLIPKAKGGEPLPEGIF 120
QY 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
DB 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
QY 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
DB 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
QY 181 FARAYAGISRTKRYMELIYEDSMDLAKLPCVAARYLRNLRYREGSGIGALDSKLSHNF 240
DB 181 FARAYAGISRTKRYMELIYEDSMDLAKLPCVAARYLRNLRYREGSGIGALDSKLSHNF 240
QY 241 TNNLGYTDHOFTELRLTYLTHSDHEGCVNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300
DB 241 TNNLGYTDHOFTELRLTYLTHSDHEGCVNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300
QY 301 HGLANOEVLVWLTOLQKEVKGVDSEKLDYIWNNTLNSGRVYPGYGHAVALRKTDPRYTQ 360
DB 301 HGLANOEVLVWLTOLQKEVKGVDSEKLDYIWNNTLNSGRVYPGYGHAVALRKTDPRYTQ 360
QY 361 REFALHLPNDPMPFKLVAQLYKIVPNVLEOGKAKMPNVDHSGVLLQYTGMTMANY 420
DB 361 REFALHLPNDPMPFKLVAQLYKIVPNVLEOGKAKMPNVDHSGVLLQYTGMTMANY 420
QY 421 TVLFGVSRALGVLAQLIWSRALGFLPERKSMSTEGIMKFEVDSKSG 466
DB 421 TVLFGVSRALGVLAQLIWSRALGFLPERKSMSTEGIMKFEVDSKSG 466

RESULT 2
ID Q9VHF5 PRELIMINARY; PRT; 466 AA.
AC Q9VHF5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE citrate synthase (EC 4.1.3.7).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Rossignol F., Jouvauille L., Mounier R., Clottes E.;
RT "Rattus norvegicus citrate synthase cDNA.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases;
CC -1- CATALYTIC ACTIVITY: CITRATE + COA -> ACETYL-COA + H(2)O +
CC OXALOACETATE.
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.
DR EMBL: AF61496; AAL6372.1; -.
DR InterPro: IPR002020; Citrate_synth.
DR Pfam: PF00285; citrate_synth.1.
DR PRINTS: PR00143; CITRATESNHASE.
DR PROSITE: PS00480; CITRATE_SYNTHASE; 1.
KW Lyase; Tricarboxylic acid cycle.
SQ SEQUENCE 466 AA; 51866 MW; AF8D38379CD38124 CRC64;

Query Match      95.4%; Score 2323; DB 11; Length 466;
Best Local Similarity 94.6%; Pred. No. 1.4e-180;
Matches 441; Conservative 15; Mismatches 10; Indels 0; Gaps 0;

QY 1 MALLTAARLLGKNNKSCVLARHNASASTNKLIDLADLPKEQARITFRQOHGKTYV 60

```

```

DB 1 MALLTAARLLGKNNKSCVLARHNASASTNKLIDLADLPKEQARITFRQOHGKTYV 60
QY 61 GOITVDMYGMKGMKGLYETSVLDPDGIIFRGFSIPECOKLIPKAKGGEPLPEGIF 120
DB 61 GOITVDMYGMKGMKGLYETSVLDPDGIIFRGFSIPECOKLIPKAKGGEPLPEGIF 120
QY 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
DB 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
QY 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
DB 121 WLLVTHIPTEEVSWLSKEMAKRAALPSHVVTMLDNFPNTHPSQLSAAVTALNSES 180
QY 181 FARAYAGISRTKRYMELIYEDSMDLAKLPCVAARYLRNLRYREGSGIGALDSKLSHNF 240
DB 181 FARAYAGISRTKRYMELIYEDSMDLAKLPCVAARYLRNLRYREGSGIGALDSKLSHNF 240
QY 241 TNNLGYTDHOFTELRLTYLTHSDHEGCVNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300
DB 241 TNNLGYTDHOFTELRLTYLTHSDHEGCVNSAHTSHLVGSALSDPYLSFAAAMNGLAGPL 300
QY 301 HGLANOEVLVWLTOLQKEVKGVDSEKLDYIWNNTLNSGRVYPGYGHAVALRKTDPRYTQ 360
DB 301 HGLANOEVLVWLTOLQKEVKGVDSEKLDYIWNNTLNSGRVYPGYGHAVALRKTDPRYTQ 360
QY 361 REFALHLPNDPMPFKLVAQLYKIVPNVLEOGKAKMPNVDHSGVLLQYTGMTMANY 420
DB 361 REFALHLPNDPMPFKLVAQLYKIVPNVLEOGKAKMPNVDHSGVLLQYTGMTMANY 420
QY 421 TVLFGVSRALGVLAQLIWSRALGFLPERKSMSTEGIMKFEVDSKSG 466
DB 421 TVLFGVSRALGVLAQLIWSRALGFLPERKSMSTEGIMKFEVDSKSG 466

RESULT 3
ID Q9CZU6 PRELIMINARY; PRT; 464 AA.
AC Q9CZU6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 2610511A05RIK protein (EC 4.1.3.7) (Citrate synthase).
GN GS OR 2610511A05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217651;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kociba H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schraml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazarrelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Sone H., Shimano H., Yamada N.;
RT "murine citrate synthase, complete.";

```



RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.  
 RL [3] SEQUENCE FROM N.A.  
 RC TISSUE-Kidney;  
 RA Strausberg R.;  
 RN Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.  
 RP [4] SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
 CC OXALOCETATE.  
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
 DR EMBL; AK012151; BAB28063.1; -  
 DR EMBL; AB056479; BAB63945.1; -  
 DR EMBL; BC013554; AAH13554.1; -  
 DR EMBL; BC029754; AAH29754.1; -  
 DR HSSP; P00889; 2CTS.  
 DR MGD; MGI:88329; CS.  
 DR InterPro; IPR002020; Citrate\_synth.  
 DR Pfam; PF00285; citrate\_synth.1.  
 DR PRINTS; PR00143; CITRATESNTHASE.  
 DR PROSITE; PS00480; CITRATE\_SYNTHASE; 1.  
 KW Lyase; Tricarboxylic acid cycle.  
 SO SEQUENCE 466 AA; 51736 MW; 710639871C31EED5 CRC64;

Query Match 95.3%; Score 2321; DB 11; Length 464;  
 Best Local Similarity 94.4%; Pred. No. 2e-180;  
 Matches 438; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY	1	MALLTAARLGTGNASCLVLAARHASASTNLKDIADLIPKEDARKITFRQOHGKTV	60
DB	1	MALLTAARLGTGNASCLVLAARHASASTNLKDIADLIPKEDARKITFRQOHGKTV	60
QY	61	GOITVDMYGGMRGKGLVETSVLPDEGIRFRGFSIPECOKLLPKAKGGEPLPEGLF	120
DB	61	GOITVDMYGGMRGKGLVETSVLPDEGIRFRGFSIPECOKLLPKAKGGEPLPEGLF	120
QY	121	WLVYGHIPTEEOVSWLSKEMAKRAALPSSHVYTMNDNPTNLHPMSQLSAAVTALNSSEN	180
DB	121	WLVYGHIPTEEOVSWLSKEMAKRAALPSSHVYTMNDNPTNLHPMSQLSAAVTALNSSEN	180
QY	181	FARAYAGISRTKRWELIYEDSMDLIAKLPVAAKIYNLREGSGIGALISNDWSHNF	240
DB	181	FARAYAGISRTKRWELIYEDSMDLIAKLPVAAKIYNLREGSGIGALISNDWSHNF	240
QY	241	TNMLGYTDHOFTELRLVLTITSHDEGGNSAHTSHLVGALSIDPYLSFAAAMNGLAGPL	300
DB	241	TNMLGYTDHOFTELRLVLTITSHDEGGNSAHTSHLVGALSIDPYLSFAAAMNGLAGPL	300
QY	301	HGLANQEVLMVLTOLQKKEVGDVSEKLDYIMNTLNSGRVVPYGAHVLKRTDPRYTCQ	360
DB	301	HGLANQEVLMVLTOLQKKEVGDVSEKLDYIMNTLNSGRVVPYGAHVLKRTDPRYTCQ	360
QY	361	REFALKHLPNDPMFLVQLKYYIVNVLLLEOGKAKNPNDASHGVLLQYGTENMY	420
DB	361	REFALKHLPNDPMFLVQLKYYIVNVLLLEOGKAKNPNDASHGVLLQYGTENMY	420
QY	421	TVLGVSRALGVLAOLINSRALGFLPRKSKSTGDKLKKFVDSK	464
DB	421	TVLGVSRALGVLAOLINSRALGFLPRKSKSTGDKLKKFVDSK	464

RESULT 4  
 Q9DAM4 PRELIMINARY; PRT; 466 AA.

AC Q9DAM4; 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE 1700007H16RIK protein (EC 4.1.3.7) (citrate synthase).  
 GN 1700007H16RIK.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21083660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schiml L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Kamuya M., Lee N. H.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kanaya M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Saeki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K. F.,  
 RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kontsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
 CC OXALOCETATE.  
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
 DR EMBL; AK005713; BAB24200.1; -  
 DR HSSP; P00889; 2CTS.  
 DR MGD; MGI:1919082; 1700007H16RIK.  
 DR InterPro; IPR002020; Citrate\_synth.  
 DR Pfam; PF00285; citrate\_synth.1.  
 DR PRINTS; PR00143; CITRATESNTHASE.  
 DR PROSITE; PS00480; CITRATE\_SYNTHASE; 1.  
 KW Lyase; Tricarboxylic acid cycle.  
 SO SEQUENCE 466 AA; 52353 MW; 10A4746A109C491C CRC64;

Query Match 91.1%; Score 2219; DB 11; Length 466;  
 Best Local Similarity 89.7%; Pred. No. 4.1e-172;  
 Matches 418; Conservative 24; Mismatches 24; Indels 0; Gaps 0;

QY	1	MALLTAARLGTGNASCLVLAARHASASTNLKDIADLIPKEDARKITFRQOHGKTV	60
DB	1	MALLTAARLGTGNASCLVLAARHASASTNLKDIADLIPKEDARKITFRQOHGKTV	60
QY	61	GOITVDMYGGMRGKGLVETSVLPDEGIRFRGFSIPECOKLLPKAKGGEPLPEGLF	120
DB	61	GOITVDMYGGMRGKGLVETSVLPDEGIRFRGFSIPECOKLLPKAKGGEPLPEGLF	120
QY	121	WLVYGHIPTEEOVSWLSKEMAKRAALPSSHVYTMNDNPTNLHPMSQLSAAVTALNSSEN	180
DB	121	WLVYGHIPTEEOVSWLSKEMAKRAALPSSHVYTMNDNPTNLHPMSQLSAAVTALNSSEN	180
QY	181	FARAYAGISRTKRWELIYEDSMDLIAKLPVAAKIYNLREGSGIGALISNDWSHNF	240
DB	181	FARAYAGISRTKRWELIYEDSMDLIAKLPVAAKIYNLREGSGIGALISNDWSHNF	240
QY	241	TNMLGYTDHOFTELRLVLTITSHDEGGNSAHTSHLVGALSIDPYLSFAAAMNGLAGPL	300
DB	241	TNMLGYTDHOFTELRLVLTITSHDEGGNSAHTSHLVGALSIDPYLSFAAAMNGLAGPL	300
QY	301	HGLANQEVLMVLTOLQKKEVGDVSEKLDYIMNTLNSGRVVPYGAHVLKRTDPRYTCQ	360
DB	301	HGLANQEVLMVLTOLQKKEVGDVSEKLDYIMNTLNSGRVVPYGAHVLKRTDPRYTCQ	360
QY	361	REFALKHLPNDPMFLVQLKYYIVNVLLLEOGKAKNPNDASHGVLLQYGTENMY	420
DB	361	REFALKHLPNDPMFLVQLKYYIVNVLLLEOGKAKNPNDASHGVLLQYGTENMY	420

OY 421 TVEGVSRALGVLAQILMSRAGLPPLERPKSMSTEGIMKFEVDSKSG 466  
 Db 421 TVEGVSRALGVLSQILMSRAGLPPLERPKSMSTEGIMKFEVDSKSG 466

RESULT 5  
 OY 421 PRELIMINARY; PRT; 464 AA.  
 ID OY401  
 AC OY401  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
 DE CG3861 protein (EC 4.1.3.7) (Citrinate synthase).  
 GN CG3861.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RX STRAIN=BERKELEY;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazet R.G., Champe M., Pfeiffer B.D.,  
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Baller R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
 RA Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Moharry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
 CC OXALOACETATE.  
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
 CC EMBL: AEO03437; AAF46159.1; -;  
 DR HSP: P00889; 2CTS;  
 DR ElyBase: Fgdn029869; CG3861.  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR Pfam: PF00285; citrate\_synth. 1.  
 DR PRINTS: PR00143; CITRATESNTASE.  
 DR PROSITE: PS00480; CITRATE\_SYNTHASE; 1.

KW Lyase: Tricarboxylic acid cycle.  
 SQ SEQUENCE 464 AA; 51574 MW; 57D582652931B4F9 CRC64;  
 Query Match 71.3%; Score 1737.5; DB 5; Length 464;  
 Best Local Similarity 70.5%; Pred. No. 6,7e-133;  
 Matches 328; Conservative 50; Mismatches 80; Indels 7; Gaps 2;

OY 1 MALLTAARLLG---TKNASCLVLAARHNASSTNIKDLADLPKQARIKTERPOOHG 56  
 Db 1 MSLYRISARKLSBAOKLPNVGAYV---RMIAADGKSLRDVLAAPQEDERVKNFRKOHG 57

OY 57 KTVVGQITVVMGGRGKMGKGLYEVSVLPDPDGIKFRGFSIPECOKLPKAKGSEPRP 116  
 Db 58 ATKMGETITDMMGKMGKIKALVETSVLDADGIRFRGLSTIPECOKVLPAAAGTEPRP 117

OY 117 EGFEMLLVTHIPIERDVSWSKEMAKRAALPSSHVVTMLDNFNTLHPNSQLSAAATYALN 176  
 Db 118 EGFEMLLVTHIPIERDVSWSKEMAKRAALPSSHVVTMLDNFNTLHPNSQLSAAATYALN 177

OY 177 SESNFAAAVAGISRTKRWELIYEDSMDLIAKLPVAAKIYRNLRYEGSGIGAIDSNDLM 236  
 Db 178 HDSKFAKAYSDGVHKSRYMEYVEDSMDLIAKLPVAAATYCTYRGKGSRSIDSLDM 237

OY 237 SHNFTNMLGTIDQFTELTLTLTSHDHEGVSASHTSHLVSALSDEPLSTFAAMNGL 296  
 Db 238 SANFVNLGIDNAPFELMLTYLTISHDEGVSASHTSHLVSALSDEPLSTFAAMNGL 297

OY 297 AGPLHGLANOEVLMTLOLKEVGVSDKLDYIMNTLNSGVNPPGSHAVLRKTPDR 356  
 Db 298 AGPLHGLANOEVLMTLOLKEVGVSDKLDYIMNTLNSGVNPPGSHAVLRKTPDR 357

OY 357 YTCQREPALKLPNDPMEKLVLAQYIVPNVLLDGGKAKNPVNDASHGVLLQYYGME 416  
 Db 358 YTCQREPALKLPNDPMEKLVLAQYIVPNVLLDGGKAKNPVNDASHGVLLQYYGME 417

OY 417 MNYTYLFGVSRAQVLAQILMSRAGLPPLERPKSMSTEGIMKFEV 461  
 Db 418 MNYTYLFGVSRAQVLAQILMSRAGLPPLERPKSMSTEGIMKFEV 462

RESULT 6  
 ID OY477 PRELIMINARY; PRT; 522 AA.  
 AC OY477;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE AT12538p (EC 4.1.3.7) (Citrinate synthase) (CG3861-PB).  
 GN CG3861.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RX STRAIN=BERKELEY;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazet R.G., Champe M., Pfeiffer B.D.,  
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Baller R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
 RA Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Moharry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
 CC OXALOACETATE.  
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
 CC EMBL: AEO03437; AAF46159.1; -;  
 DR HSP: P00889; 2CTS;  
 DR ElyBase: Fgdn029869; CG3861.  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR Pfam: PF00285; citrate\_synth. 1.  
 DR PRINTS: PR00143; CITRATESNTASE.  
 DR PROSITE: PS00480; CITRATE\_SYNTHASE; 1.

RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Bokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jaitali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Mishina N.V., Modarity C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri I., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Sidenklamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Welschenbach J.,  
 RA Williams S.M., Woodger, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [3]

RN SEQUENCE FROM N.A.  
 RP Ceiniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Cocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Barzon J.J., An H., Baldwin D., Barzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett E., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jaitali M., Kruse D., Li P., Mettel B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,  
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri I., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [4]

RN SEQUENCE FROM N.A.  
 RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochownik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Bernan B., Carlson J.M., Ceiniker S.E.,  
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [5]

RN SEQUENCE FROM N.A.  
 RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
 CC OXALOACETATE.  
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
 CC EMBL: AY089318; AAL90056.1; -  
 DR EMBL: AE003437; AAN09169.1; -

DR FlyBase: FBgn0029869; CG3861.  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR Pfam: PF00285; citrate\_synth. 1.  
 DR PROSITE: PS00480; CITRATE\_SYNTHASE; 1.  
 KW Lyase; tricarboxylic acid cycle.  
 SO SEQUENCE 522 AA; 58217 MW; 7CA9D79350F2764D CRC64;

Query Match 71.2% Score 1735; DB 5; Length 522;  
 Best Local Similarity 73.3%; Pred. No. 1.3e-132;  
 Matches 321; Conservative 48; Mismatches 69; Indels 0; Gaps 0;

OY 24 RHASASTNLKDLADLIPKEDARIKTFROQHGTVVGQITVDMYGMGMKGLVYETS 83  
 DB 83 RMAADGSLRDVLAAPQOQERVKFRKOGATKMGETTIDMVGMRGKALVYETS 142  
 OY 84 VLDPDEGRFRFGFSIPKQKLLPRAKGEERLPGLFVLLVTHGHPPEQVSMKSKPAK 143  
 DB 143 VLADDEGIRFRLGSLIPKQKLLPRAKGEERLPGLFVLLVTHGHPPEQVSMKSKPAK 202  
 OY 144 RAALPSHVVTMLDNEPTLHPMSOLSAVTLANSESNFARAYAGISRTKYMELIEDSM 203  
 DB 203 RAALPQHVTMLNMPPTLHPMSOPAAVTLANSDSKAKAYSDGVKSKYWEYEDSM 262  
 OY 204 DLIAKLPVAAKIRYRNLREGSGIGALDSNDSNHNFTNMLGYTDHOFTELRLYLTHS 263  
 DB 263 DLIAKLPVAAKIRYRNLREGSGIGALDSNDSNHNFTNMLGYTDHOFTELRLYLTHS 322  
 OY 264 DHEGNSVAHSHLVGSLSPDYLSFAAAMNGLGRLHGLNOLVWLTOLQREVGNDV 323  
 DB 323 DHEGNSVAHSHLVGSLSPDYLSFAAAMNGLGRLHGLNOLVWLTOLQREVGNDV 382  
 OY 324 SDEKLDVYTWNTLNSGRVPGYGHAVLKRTPRYTCOREFALKHLPNPMKLYAQLTKI 383  
 DB 383 SEEDLKEVYTWNTLNSGRVPGYGHAVLKRTPRYTCOREFALKHLPNPMKLYAQLTKI 442  
 OY 384 VPVNLLEGGKAKNPVNDASHVLLQYTGTEKNYTVLFGVSALGVLAQLIWSRALG 443  
 DB 443 VPVNLLEGGKAKNPVNDASHVLLQYTGTEKNYTVLFGVSALGVLAQLIWSRALG 502  
 OY 444 FPLEPRKSMSTEGMLKPFV 461  
 DB 503 LPIERPSFSTDLVKMV 520

RESULT 7  
 ID 08NKF2 PRELIMINARY; PRT; 474 AA.  
 AC 08NKF2  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 GN Citrate synthase (EC 4.1.3.7).  
 OS *Emmericella nidulans* (*Aspergillus nidulans*).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 CC *Emmericella*; Trichocomaceae; *Emmericella*.  
 OX NCBI\_TaxID=162425;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-FGSC4;  
 RA Seo S.W., Lee C.H., Maeng P.J.,  
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
 CC OXALOACETATE.  
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
 CC EMBL: AF468824; AAM22645.1; -  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR Pfam: PF00285; citrate\_synth. 1.  
 DR PRINTS: PR00143; CITRATESYNTHASE.  
 DR PROSITE: PS00480; CITRATE\_SYNTHASE; 1.  
 KW Lyase; tricarboxylic acid cycle.  
 SO SEQUENCE 474 AA; 52160 MW; 879F9E95861ABDCF CRC64;

Query Match	63.6%	Score 1349	DB 3	Length 474
Best Local Similarity	65.3%	Pred. No. 1.5e-117		
Matches	288	Conservative	67	Indels 2
				Gaps 2
QY	25	HASASSTMLKDIALDILPEKQARIRKTFPOQHGTQVVGQTVVMYMGHMGKMGLEYENSV	84	
DB	34	YSTGTRKSLKEFPAKLPGLKELEKVKLKEHNNKVIIGELTLDQATGARGVCLYMEGSV	93	
QY	85	LDPDEGIRFRCFSIDECCOKLLPKRANGEBPLREGFLFWLLVTGHIPTEDOVSWLSKEMAKR	144	
DB	94	LDSEEGIRFRGLTPECCOKLLPKRANGEBPLREGFLFWLLVTGHIPTEDOVSWLSKEMAKR	153	
QY	145	AALPEHYVTMLDNFPTNLHPMSQSLAAATLANSESNFAPRAYNOGTSRTKYWELIYEDSMD	204	
DB	154	SDLPEFTIEBLDRVPSTLHPAQESFLAVATLHEHSAFAAKVGLINKKEVHYHTEDSMD	213	
QY	205	LIAPKPCVAAKIYRMLYREGSGIGADISNLMDSHNFNTMLGTYDHO-FTELRLYLTIHS	263	
DB	214	LIAPKPTIAAKIYRNVFKDK-VAPIDKDKDYSYLANQDGLPADKKDFEIMRLVLTIS	272	
QY	264	DHEGNNVAHSHLYGALSIDPYLSFAAAMGCLAPLHGLANQEVLMITLOKEVGDV	323	
DB	273	DHEGNNVAHSHLYGALSIDPYLSFAAAMGCLAPLHGLANQEVLMITLOKEVGDV	332	
QY	324	SDEKLRDYIMNTLNSGRVVPVGGHVLNRTDPRYTCOREFALKHLPNDPEKLVNOLYKI	383	
DB	333	SDQSLKIDMLSTLNGRVVPVGGHVLNRTDPRYTCOREFALKHLPNDPEKLVNOLYKI	392	
QY	384	VPNVLLGOGKAKPNPNVDANHSGLVLYOYGMTENMYVYVLFQVSRLAGVLAQLIMSRLG	443	
DB	393	APGVLTGHTKTPNPVNDANHSGLVLYOYGMTENMYVYVLFQVSRLAGVLAQLIMSRLG	452	
QY	444	FLERPKSMTSEGLMKFVDSK	464	
DB	453	APIERPKSFSTEAVAKLVGAK	473	
RESULT 8				
ID	Q9P8D6	PRELIMINARY:	PRT:	474 AA.
AC	Q9P8D6:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DR	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DE	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
GN	Citrate synthase precursor (EC 4.1.3.7).			
OS	Aspergillus niger.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiaceae; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.			
NCBI	NCBI_TaxID=5061;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=N400;			
RA	MEDLINE=2015633; PubMed=10689162;			
RA	Ruijter G.J., Panneman H., Xu D.B., Visser J.;			
RT	"Properties of Aspergillus niger citrate synthase and effects of citra			
RT	overexpression on citric acid production.";			
RL	FEMS Microbiol. Lett. 184:35-40(2000).			
CC	-1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2O) +			
CC	OXALOACETATE.			
CC	-1- PATHWAY: TRICARBOXYLIC ACID CYCLE.			
CC	-1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.			
DR	EMBL: AJ243204; CAB7625.1; -			
DR	HSSP: P23007; 2CSG.			
DR	InterPro: IPR002020; Citrate_synth.			
DR	Pfam: PF00285; citrate_synth; 1.			
DR	PRINTS: PR00143; CITRYSNTHASE.			
DR	PROSITE: PS00480; CITRATE SYNTHASE; 1.			
KN	Lysase; Signal: Tricarboxylic acid cycle.			
FT	CHAIN 1 34			
FT	POTENTIAL.			
FT	POTENTIAL.			
SEQUENCE	474 AA: 52025 MW; 62646ABF196994A1D0 CRC64;			

```

Query Match      63.2%, Score 1539.5, DB 3, Length 474;
Best Local Similarity 61.9%, Pred. No. 9.2e-117;
Matches 293; Conservative 69; Mismatches 96; Indels 15; Gaps 3;

OY 6 AARRLTGRKNASCVLVAR-----HASASTNLKDLADLIPRECARITFR 52
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2 ASTLRLTGSLARSTSIARKPVVQSAATFNGRLRCYTGAKSLIKETFAEKLPAEIKYVKLR 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 53 QOHGKTVVQGITVDMYTGGMGRMGKGLVYETSLDPRDEGIRPFGFSIEPCOKLLPKARAGE 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 62 KEHSKSVIGEVTLDQAGVAGARGVAKCLWEGSVLSEDEGIRPGRRTIEPCQELLPKARGQ 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 113 EPLDEGLEFWLLVTHGHIPEDEQVSLSKEMAKRALPESHVYTMIDNPTNLHPMSQLSAAV 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 122 EPLDEGLEFWLLLTLEIPEDQVRLSSEMAKRSLPFIEELIDRCPSTLHPMSQSFVAV 181
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 173 TALNSESNFARAVAQGISRTKYWELIYEDSDMLDAKLPICVAKIYRNLRYEGSGIGADIS 232
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 182 TALHESAFKAAYAKGINKKQYMYVTFEDSDMLDAKLPITAAKIYRNVFNDGK -VARIQ 240
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 233 NLDMSHNFNTMAGYTD -HQFTELRLVLTTHSDHEGNGVSAHSHLVGALSIDPYLSFAA 291
   : : : : : : | : | : | : | : | : | : | : | : | : | : | : |
DB 241 DKDYSYNIANOLGCGDNDNFVELMRRLVLTTHSDHEGNGVSAHSHLVGALSISPMLSIAA 300
   : : : : : : | : | : | : | : | : | : | : | : | : | : | : |
OY 292 AMNGIAGLGHLANOEYVLTWLTOLQKEKGVDSXDEKLRDITWNTLNSGRVVPYGGHAVLR 351
   : : : : : : | | | | | | | | | | | | | | | | | | | | | |
DB 301 GLNGIAGLGHLANOEYVLTWLTOLQKKAALAGNLSDEALIKNTYMTSLNAGVYPGGHAVLR 360
   : : : : : : | | | | | | | | | | | | | | | | | | | | | |
OY 352 KTDPRYTCQRFALKAHPNDPMFPLVQLYKIVPNVLLEOGKANPMPNVDAHSGVLLQY 411
   : : : : : : | : | : | : | : | : | : | : | : | : | : | : |
DB 361 KTDPRYVSQREFALRKLPDDPMFLQSVYKVIAGVLTLEHGKTKNPYPNVDAHSGVLLQY 420
   : : : : : : | : | : | : | : | : | : | : | : | : | : | : |
OY 412 YGKTEMYVTVLFEVSRALGYLAOLINSRALGFLERPKSKSTGKLMKPFVDSK 464
   | : | : | | | | | | | | | | | | | | | | | | | | | | | |
DB 421 YGLEANYVTVLFEVSRALGYLPOLLIDRALGADIERPKSYSTEAFKLVGAK 473
   | : | : | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
O9HF11 PRELIMINARY; PRT; 472 AA.
AC O9HF11;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DEF Mitochondrial citrate synthase (EC 4.1.3.7).
GN CN.
OS Podospora anserina.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Lasiosphaeriaceae; Podospora.
OC NCBI_TaxID=5145;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S;
RC Rupplich-Robert G., Zickler D., Berteaux-Lecellier V., Picard M.;
RT "Mutations in Podospora citli gene act as partial suppressors of a
RT peroxisomal mutant and impair progression of meiosis downstream of the
RT diffuse stage."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +
CC OXALACETATE.
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.
CC EMBL; AJ296102; CAC12961.1; -.
DR HSSP; P23007; 2CSC.
DR InterPro; IPR002020; Citrate-synt.
DR Pfam; PF00285; citrate-synt; 1.
DR PRINTS; PR00143; CITRTSNTHASE.
DR PROSITE; PS00480; CITRATE_SYNTHASE; 1.
KW lase; Tricarboxylic acid cycle.
SEQUENCE 472 AA: 52309 MW; 44608347095D8D91 CMC64;

Query Match      61.9%, Score 1507; DB 3; Length 472;

```

	Matches	276	Conservative	72	Mismatches	91	Indels	3	Gaps	3
22	AARHASSYNNLKDLADLPKEQARIKTFQHQGKTVVGOITVDMYGGMRGMGLAYE	81								
OY		11	:   :	11	: :	1	: :	1		1

```

Db 10 ATAAVTAETLTKEERAEALLPGQAEKELSKHGKTVIGEVLLLEQAVGMRGKIGLWME 69
QY 82 TSVLDBDEGIRFGFSIPECQKLLPKAKGEEPLPGELFWLLVTGHIPTREEQVSWLSKEW 141
Db 70 GSVLDPVEGIRFGKTIPIQKELPKAKGSESEPLPEALLFTLGTVEPLTAQTKAASAE 129
QY 142 AKKAALPSHVYVTLNDPEPTLHPMSQLSAAVTLNSESFNARAYAGISITKWTWELIYED 201
Db 130 ASRSALPKHVEDLDLPCPTLHPMAQFSAINALESESKAKAVAGVNNKEKWKYTED 189
QY 202 SMDLKLKPCVAAKIRNLYREGSGIGADISNLDWMSHNFNMIGYTDH-OFTELTRLTYLT 260
Db 190 SIDLLAKLPVAAKIRNRYFKDCK-VAPVNASLDYSHNFANMLGFGDNKFEVFLMRLTYS 248
QY 261 IHSDEGNGVSAHTSHLVGSALSDPYLSFAAAMNGLAGPLGLANGEVLMWLTQLOKEVG 320
Db 249 IHSDEGNGVSAHTSHLVGSALSDPYLSFAAAMNGLAGPLGLANGEVLMWLTQLOKEVG 308
QY 321 KQV-SDEKLRDYIWNVLNSGRVYPGCHAVLRTKTDPRYTCQREFALKHLPENDMEKLYAQ 379
Db 309 DDIENKEKIEEYLMKTLNAGRVYGVHAVALRTDPRYTAQREFALKHMDYDMFKLYSN 368
QY 380 LYKIVPNVLLQOGKAKNPNPNVDAHSGVLLQYGYGMEEMNYTVLFGVSALGVLAOLIMS 439
Db 369 IYEVAKRYLAEOGKTNNPNPNVDSHSGIILQYGYGLQESYTYVLFQVAAAFVLPOLLID 428
QY 440 RALGFLPERPKSMSTEGLMKFVDS 463
Db 429 RGLGMAIERPKSFSTQKYELEN 452

```

## RESULT 12

```

ID 08SSP7 PRELIMINARY; PRT; 460 AA.
AC 08SSP7;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Putative lyase (EC 4.1.3.7) (Citrate synthase).
OS Dicyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dicyostelida; Dicyostelium.
OX NCBI_Taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AT4;
RA Goeckner G., Eichinger L., Szafiranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Lehmann B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dicyostelium.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +
CC OXALOACETATE.
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.
DR EMBL: AC116960; AA08461.1; -.
DR InterPro: IPR002020; Citrate_synth.
DR Pfam: PF00285; citrate_synth; 1.
DR PROSITE: PS00480; CITRATE_SYNTHASE; 1.
DR Lysase; Tricarboxylic acid cycle.
SQ SEQUENCE 460 AA; 51220 MW; 1B4A0B58E4519E0 CRC64;

```

Query Match 60.9%; Score 1483.5; DB 5; Length 460;  
 Best Local Similarity 61.7%; Pred. No. 3.2e-112;  
 Matches 282; Conservative 66; Mismatches 104; Indels 5; Gaps 3;

```

QY 8 ARLLGTKNASCLVLAARHASSTNLKDLIADLPKQARITKFRPOHGKTVYGOITVDM 67
Db 3 ARLLGTKNASCLVLAARHASSTNLKDLIADLPKQARITKFRPOHGKTVYGOITVDM 67
QY 68 MYGGMGMGLVETSVLDBDEGIRFGFSIPECQKLLPKAKGEEPLPGELFWLLVTGHI 127
Db 59 AYGMSVSKSLVETSVLDBDEGIRFGFSIPECQKLLPKAKGEEPLPGELFWLLVTGHI 118

```

```

QY 128 IPTREEVSWLSKEMAKRAALPSHVYVTLNDPEPTLHPMSQLSAAVTLNSESFNARAYAQ 187
Db 119 VPTREOVKTLSEKLAARAGIIPKHTVSMIRAFPEQPMQSLAAAIATLQGESKFEVAYND 178
QY 188 GISRTKWTWELIYEDSDSLDAKLPEVAAKIRNLYREGSGIGADISNLDWMSHNFNMIGY 247
Db 179 GAKDKYMSSTLEDSLDVAKLPEVAAKIRNLYREGSGIGADISNLDWMSHNFNMIGY 238
QY 248 DHOFTELTRLTYLTSHSDEGNGVSAHTSHLVGSALSDPYLSFAAAMNGLAGPLGLANGE 307
Db 239 SKQFDELMRLTYLTSHSDEGNGVSAHTSHLVGSALSDPYLSFAAAMNGLAGPLGLANGE 298
QY 308 VLVMTQLOKEVG-KQVSDKLRDYIWNVLNSGRVYPGCHAVLRTKTDPRYTCQREFALK 366
Db 299 VLVMTQLOKEVG-KQVSDKLRDYIWNVLNSGRVYPGCHAVLRTKTDPRYTCQREFALK 358
QY 367 HLPNDMEFKLYAQVLYKIVPNVLLQOGKAKNPNPNVDAHSGVLLQYGYGMEEMNYTVLFGV 426
Db 359 HLPNDMEFKLYAQVLYKIVPNVLLQOGKAKNPNPNVDAHSGVLLQYGYGMEEMNYTVLFGV 418
QY 427 SRALGVLAOLIMSRLGFLPERPKSMSTEGLMKFVDS 463
Db 419 SRALGVLAOLIMSRLGFLPERPKSMSTEGLMKFVDS 455

```

## RESULT 13

```

ID 09FUJ7 PRELIMINARY; PRT; 472 AA.
AC 09FUJ7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Citrate synthase (EC 4.1.3.7).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriophytaceae; Oryzaeae; Oryza.
OX NCBI_Taxid=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Scival da Silva A.L., Becker D., Loerz H.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +
CC OXALOACETATE.
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.
DR EMBL: AE302906; AA62877.1; -.
DR HSSP: P23007; 2CSC.
DR Gramene: 09FUJ7; -.
DR InterPro: IPR002020; Citrate_synth.
DR Pfam: PF00285; citrate_synth; 1.
DR PRINTS: PR00143; CITRATESYNTHASE.
DR PROSITE: PS00480; CITRATE_SYNTHASE; 1.
DR Lysase; Tricarboxylic acid cycle.
SQ SEQUENCE 472 AA; 52228 MW; 52F7A2003EFE1811 CRC64;

```

Query Match 60.8%; Score 1481.5; DB 10; Length 472;  
 Best Local Similarity 60.5%; Pred. No. 4.8e-112;  
 Matches 277; Conservative 80; Mismatches 96; Indels 5; Gaps 3;

```

QY 4 LTAARLGS--TKNASCL--VLAARHASSTNLKDLIADLPKQARITKFRPOHGKTVY 59
Db 7 LTAARLGS--TKNASCL--VLAARHASSTNLKDLIADLPKQARITKFRPOHGKTVY 59
QY 60 VGOITVDMYMGARKGKGLVETSVLDBDEGIRFGFSIPECQKLLPKAKGEEPLPGEL 119
Db 67 LGMITVDMYMGARKGKGLVETSVLDBDEGIRFGFSIPECQKLLPKAKGEEPLPGEL 126
QY 120 FWLLVTGHIPTREEVSWLSKEMAKRAALPSHVYVTLNDPEPTLHPMSQLSAAVTLNSES 179
Db 127 LWMLLTGKVPKTEQVVALSKELASRSSGVGHVKAIDALPVTAHPMTQFTTGVMALQVES 186
QY 180 NFARAYAGISITKWTWELIYEDSDSLDAKLPEVAAKIRNLYREGSGIGADISNLDWMSH 239

```

Db 187 EFQAYDKGMSKSFWEPTEDCNLTARPAVASYYRRJFKGKTIAA-DNALDYAAN 245  
Qy 240 FTNNLGYTDHFTLRLYLTTHSDHSGNVAHTSHLVGSLADPYLSFAAANGLAGP 299  
Db 246 FSHMLGDDPKMLMLRLYTHIDHSGNVAHTSHLVGSLADPYLSFAAANGLAGP 305  
Qy 300 LHGLANDEVLMITOLQKENVGKDVSDKLRDYINNTJNSGRVPGYGHAVLRKTDPRYTC 359  
Db 306 LHGLANDEVLMITOLQKENVGKDVSDKLRDYINNTJNSGRVPGYGHAVLRKTDPRYTC 365  
Qy 360 QREPALKLPNDPKFVLAQVLYKIVPNVLEQGAKKPNPNDVDAHSGVLLQYGMTEKNY 419  
Db 366 QREPALKLPNDPKFVLAQVLYKIVPNVLEQGAKKPNPNDVDAHSGVLLQYGMTEKNY 425  
Qy 420 YTVFEGVSRALGVAOLIMSRALGFPLERPKSMTEGL 457  
Db 426 YTVFEGVSRALGVAOLIMSRALGFPLERPKSMTEGL 463

## RESULT 14

Qy 096544 PRELIMINARY: PRT: 437 AA.  
AC 096544;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Citrate (Si)-synthase (EC 4.1.3.7) (Citrate synthase)  
DE (Fragment).  
CN CITI.  
OS Beta vulgaris (Sugar beet).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Beta.  
OX NCBI\_TaxID=161934;  
RN [1]  
RP SEQUENCE FROM N.A.  
RL La Cognata U., Landschuetze V., Willmitzer L., Mueller-Roeber B.;  
RA Plant Cell Physiol. 0:0-0(0).  
CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
CC OXALACETATE.  
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
DR EMBL: X84228; CAAS9010.1; -  
DR HSSP: P23007; 2CSC.  
DR InterPro: IPR002020; Citrate\_synth.  
DR Pfam: PF00285; citrate\_synth.  
DR PRINTS: PR00143; CITRISNTHASE.  
DR Lyase; Tricarboxylic acid cycle.  
KW NON\_TER 1  
SQ SEQUENCE 437 AA; 48881 MW; FB50CA21ED5519A3 CRC64;

Query Match 60.8%; Score 1480.5; DB 10; Length 437;  
Best Local Similarity 62.6%; Pred. No. 5.2e-112;  
Matches 271; Conservative 70; Mismatches 91; Indels 1; Gaps 1;

Qy 28 ASSTNLADLADLPKQOARIKTPROOQKTVGQITVDMMYGMGMKGLVETSYDLP 87  
Db 1 SSNDLASELOELIPBOOERAKTKKEFGSLQNLINVDVLCGMKRGHTGLMETSLDLP 60  
Qy .88 DEGIRFGFSIPECOKLLPKAKGGEPLPEGLFWLVLTGHIPTTEOVWSLSEMAKRAAL 147  
Db 61 EEGIRFGFSIPECOKLLPKAKGGEPLPEGLFWLVLTGHIPTTEOVWSLSEMAKRAAL 120  
Qy 148 PSHVVTMLDNPPTLHMPMSOLSAVTLANSNSFARAYAGISRTKRWELIYEDSMOLA 207  
Db 121 PDHYKTIIDALPTAHPTQCTGVMLQTFSEFOKAYEKGIHNSKFWPEYEDCLSLIA 180  
Qy 208 KLPCVAAKIYRNLTREGSGAIDSNLDMSHNFTNMLGYTDHFTLRLYLTTHSDHSG 267  
Db 181 QVPVAAVYVRRMKNGQVI-PLDDSLIDYGGNFAMHGLCFDPSQMLELRKLVYTHSDHSG 239  
Qy 268 GNSAHTSHLVGSLADPYLSFAAANGLAGPLHGLANDEVLMITOLQKENVGKDVSDK 327

Db 240 GNSAHTSHLVGSLADPYLSFAAANGLAGPLHGLANDEVLMITOLQKENVGKDVSDK 299  
Qy 328 LRDYIWNLTNSGRVPGYGHAVLRKTDPRYTCOREFALKHLPNDPKFVLAQVLYKIVPNV 387  
Db 300 LKDYVMTLNSGKVPGGGLVLRKTDPRYTCOREFALKHLPNDPKFVLAQVLYKIVPNV 359  
Qy 388 LLEQGAKKPNPNDVDAHSGVLLQYGMTEKNYVTVLFGVSRALGVAOLIMSRALGFPL 447  
Db 360 LLEQGAKKPNPNDVDAHSGVLLQYGMTEKNYVTVLFGVSRALGVAOLIMSRALGFPL 419  
Qy 448 RPKSMTEGLMKF 460  
Db 420 RPKSMTEGLMKF 432

## RESULT 15

Qy 09P4E3 PRELIMINARY: PRT: 475 AA.  
AC 09P4E3;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Putative citrate synthase (EC 4.1.3.7).  
OS Saccharomyces kluyveri (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
OX NCBI\_TaxID=4934;  
RN [1]  
RP SEQUENCE FROM N.A.  
RL STRAIN-IFO 1894;  
RA Langkjaer R.B., Nielsen M.L., Dagaard P.R., Liu W., Piskur J.;  
RT "Yeast chromosomes have been significantly reshaped during their  
evolutionary history."  
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
CC OXALACETATE.  
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
DR EMBL: AF193854; AAF78896.1; -  
DR HSSP: P23007; 2CSC.  
DR InterPro: IPR002020; Citrate\_synth.  
DR Pfam: PF00285; citrate\_synth.  
DR PRINTS: PR00143; CITRISNTHASE.  
DR PROSITE: PS00480; CITRATE SYNTHASE; 1.  
KW Lyase; Tricarboxylic acid cycle  
SQ SEQUENCE 475 AA; 52865 MW; E49406B2C4729F4 CRC64;

Query Match 60.8%; Score 1480.5; DB 3; Length 475;  
Best Local Similarity 60.8%; Pred. No. 5.9e-112;  
Matches 276; Conservative 71; Mismatches 106; Indels 1; Gaps 1;

Qy 10 LIGTKNASCVLARHNASASTNLKOLLADLPKQOARIKTPROOQKTVGQITVDMMY 69  
Db 18 LKSSKTGGLFGLAARNYSNGEKTLEKRFALFESKAEIKLKLKHEKQTYIGEVLLSOAY 77  
Qy 70 GGMGMKGLVETSYDLPDEGIRFGFSIPECOKLLPKAKGGEPLPEGLFWLVLTGHIPT 129  
Db 78 GGMGMKGLVETSYDLPDEGIRFGFSIPECOKLLPKAKGGEPLPEGLFWLVLTGHIPT 137  
Qy 130 TEEQVWSLSEMAKRAALPSHVVTMLDNPPTLHMPMSOLSAVTLANSNSFARAYAGI 189  
Db 138 TEEQVWSLSEMAKRAALPSHVVTMLDNPPTLHMPMSOLSAVTLANSNSFARAYAGI 197  
Qy 190 SRTYVWELIYEDSMOLAPKQOARIKTPROOQKTVGQITVDMMYGMGMKGLVETSYDLP 249  
Db 198 AKKQYVWELIYEDSMOLAPKQOARIKTPROOQKTVGQITVDMMYGMGMKGLVETSYDLP 256  
Qy 250 QFTLRLYLTTHSDHSGNVAHTSHLVGSLADPYLSFAAANGLAGPLHGLANDEVLM 309  
Db 257 DFVDMMLYLTTHSDHSGNVAHTSHLVGSLADPYLSFAAANGLAGPLHGLANDEVLM 316  
Qy 310 VMLTOLQKENVGKDVSDKLRDYINNTJNSGRVPGYGHAVLRKTDPRYTCOREFALKHLP 369







Query Match 98.5%; Score 2399; DB 1; Length 466;  
 Best Local Similarly 98.7%; Pred. No. 7.2e-173;  
 Matches 460; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 MALTTAARLLGTRKNSCLVLAARHNASSTNLKDLADLIPKROARIKTRPOOHGTVV 60  
 1 MALTTAARLLGTRKNSCLVLAARHNASSTNLKDLADLIPKROARIKTRPOOHGTVV 60  
 61 GOITVDMVMYGMGMKGLVETSYLDPDEGIRFEGSFISPCOKLLPRAKGEEPLPEGLF 120  
 61 GOITVDMVMYGMGMKGLVETSYLDPDEGIRFEGSFISPCOKLLPRAKGEEPLPEGLF 120  
 61 GOITVDMVMYGMGMKGLVETSYLDPDEGIRFEGSFISPCOKLLPRAKGEEPLPEGLF 120  
 121 WLVTGCIPTPEEYQSWLSKMAKRALPSHVVTMLDNFNNLPHMSQLSAAYVALNSES 180  
 121 WLVTGCIPTPEEYQSWLSKMAKRALPSHVVTMLDNFNNLPHMSQLSAAYVALNSES 180  
 181 FARAYAGISRTKYWEIYEDSDMLAKPCVAKIYRNLREGSGIGALDSNLDMSHNF 240  
 181 FARAYAGISRTKYWEIYEDSDMLAKPCVAKIYRNLREGSGIGALDSNLDMSHNF 240  
 181 FARAYAGISRTKYWEIYEDSDMLAKPCVAKIYRNLREGSGIGALDSNLDMSHNF 240  
 241 TNNLGYTDHOFTELTLRLYLTHSDHEGCVSAHTSHLVGSALSDPYLSFAAAMGLAGPL 300  
 241 TNNLGYTDHOFTELTLRLYLTHSDHEGCVSAHTSHLVGSALSDPYLSFAAAMGLAGPL 300  
 301 HGLANOEVLWVLQLOKEVEKDVSDKLRDYINNTLNSGRVPGYGAVALRKTDPRYTCQ 360  
 301 HGLANOEVLWVLQLOKEVEKDVSDKLRDYINNTLNSGRVPGYGAVALRKTDPRYTCQ 360  
 361 REFALKHLPNDPMPKLAQLYKYIYPNVLLFOGKAKNPVNDASHGVLQYGGTEENY 420  
 361 REFALKHLPNDPMPKLAQLYKYIYPNVLLFOGKAKNPVNDASHGVLQYGGTEENY 420  
 421 TVLFGVSRAVGVALQILWSRALGFLPRKSMSTEGIMKRVDSKSG 466  
 421 TVLFGVSRAVGVALQILWSRALGFLPRKSMSTEGIMKRVDSKSG 466

RESULT 2  
 CIST\_PIG STANDARD; PRT; 464 AA.  
 ID CIST\_PIG STANDARD; PRT; 464 AA.  
 AC P00889;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Citrate synthase, mitochondrial precursor (EC 2.3.3.1).  
 GN CS.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID-9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89000665; PubMed=3048387;  
 RA Evans C.T., Owens D.D., Sumegi B., Kispal G., Sreere P.A.;  
 RT "Isolation, nucleotide sequence, and expression of a cDNA encoding  
 pig citrate synthase.";  
 RT Biochemistry 27:4680-4686(1988).  
 RN [2]  
 RP SEQUENCE OF 28-464.  
 RX MEDLINE=8223193; PubMed=7093227;  
 RA Bloxham D.P., Parmelee D.C., Kumar S., Walsh K.A., Titani K.;  
 RT "Complete amino acid sequence of porcine heart citrate synthase.";  
 RT Biochemistry 21:2028-2036(1982).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 AND 1.7 ANGSTROMS).  
 RX MEDLINE=83010291; PubMed=7120407;  
 RA Remington S., Wiegand G., Huber R.;  
 RT "Crystallographic refinement and atomic models of two different forms  
 of citrate synthase at 2.7- and 1.7-A resolution.";  
 RT J. Mol. Biol. 158:111-152(1982).  
 RN [4]  
 RP MUTAGENESIS.  
 RX MEDLINE=91104711; PubMed=1702991;

RA Alter G.M., Casazza J.P., Zhi W., Nemeth P., Sreere P.A., Evans C.T.;  
 RT "Mutation of essential catalytic residues in pig citrate synthase.";  
 RL Biochemistry 29:7557-7563(1990).  
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate = citrate +  
 COA.  
 CC -1- PATHWAY: Tricarboxylic acid cycle.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS  
 CAPABLE OF OXIDATIVE METABOLISM.  
 CC -1- SIMILARITY: Belongs to the citrate synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M21197; AAA31017.1; -.  
 DR PIR: A29966; YKRG.  
 DR PDB: 1CTS; 16-JUL-88.  
 DR PDB: 2CTS; 09-OCT-88.  
 DR PDB: 3CTS; 09-OCT-88.  
 DR PDB: 4CTS; 16-JUL-88.  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR Pfam: PF00285; citrate\_synth.  
 DR PRINTS: PR00143; CITRSMTHASE.  
 DR PROSITE: PS00480; CITRATE SYNTHASE; 1.  
 DR Transiterase: Tricarboxylic acid cycle; Mitochondrion; 3D-structure;  
 KM Transist peptide: Methylation.  
 KW TRANSIT 1  
 FT CHAIN 28 464  
 FT MOD RES 395 395  
 FT ACT\_SITE 301 301  
 FT ACT\_SITE 347 347  
 FT ACT\_SITE 402 402  
 FT HELIX 33 55  
 FT TURN 56 57  
 FT TURN 65 69  
 FT HELIX 70 71  
 FT TURN 73 74  
 FT STRAND 84 86  
 FT TURN 87 89  
 FT STRAND 90 92  
 FT STRAND 97 97  
 FT HELIX 98 104  
 FT STRAND 107 107  
 FT STRAND 114 114  
 FT HELIX 116 125  
 FT TURN 131 143  
 FT TURN 144 144  
 FT HELIX 149 157  
 FT HELIX 164 174  
 FT HELIX 175 178  
 FT HELIX 180 188  
 FT TURN 187 188  
 FT TURN 191 193  
 FT HELIX 194 221  
 FT TURN 223 224  
 FT TURN 232 233  
 FT HELIX 236 244  
 FT TURN 245 245  
 FT TURN 249 261  
 FT TURN 262 262  
 FT TURN 266 268  
 FT TURN 270 279  
 FT TURN 280 282  
 FT HELIX 285 296  
 FT TURN 297 297  
 FT TURN 299 303  
 FT HELIX 304 318



FT TURN 218 218  
 FT HELIX 222 234  
 FT TURN 235 235  
 FT HELIX 243 252  
 FT TURN 253 255  
 FT HELIX 258 269  
 FT TURN 270 270  
 FT TURN 272 276  
 FT HELIX 277 291  
 FT TURN 292 293  
 FT HELIX 298 310  
 FT TURN 311 312  
 FT STRAND 316 317  
 FT TURN 318 318  
 FT HELIX 328 340  
 FT TURN 342 343  
 FT HELIX 345 364  
 FT TURN 365 365  
 FT STRAND 372 372  
 FT TURN 374 374  
 FT HELIX 375 384  
 FT TURN 385 386  
 FT HELIX 390 392  
 FT HELIX 393 414  
 FT TURN 415 416  
 FT HELIX 427 433  
 SQ SEQUENCE 433 AA: 47373 MW: 6942294BA95A9E06 CRC64;

Query Match 84.1%; Score 2048; DB 1; Length 433;  
 Best Local Similarity 90.5%; Pred. No. 1.6e-146;  
 Matches 390; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

QY 28 ASSTNKDLIADLIPKQARIKTFROOHGKTIVVQITVDMYGMGMKGLVETSYLDP 87  
 DB 1 ASSTNKDVYLAALIPKQARIKTFROOHGKTALGOITVDMSYGMGMKGLVETSYLDP 60  
 QY 88 DEGIRRGFSIPKQKLLPRKAGGEPLEPGLEFVLVTGHIPIPEEOYSMLSKEMAKRAAL 147  
 DB 61 DEGIRRGFSIPKQKLLPRKAGGEPLEPGLEFVLVTGHIPIPEEOYSMLSKEMAKRAAL 120  
 QY 148 PSHVVTMLDNFPTNLHPMSQLSAVYALNSENSENFARAAYAGISRTKYMWELIVEDSMOLA 207  
 DB 121 PSHVVTMLDNFPTNLHPMSQLSAVYALNSENSENFARAAYAGISRTKYMWELIVEDSMOLA 180  
 QY 208 KLPCVAAKTYRNLIRBEGSIGAIDSNLDMSHNTNMLGYTDHOFTELRLYLTIHSHDEG 267  
 DB 181 KLPCVAAKTYRNLIRBEGSIGAIDSNLDMSHNTNMLGYTDHOFTELRLYLTIHSHDEG 240  
 QY 268 GNVSAHTSHLVGSALSDPYLFAAAMNGLAGPLHGLANOEVLMLOLOKEVGKDYSDK 327  
 DB 241 GNVSAHTSHLVGSALSDPYLFAAAMNGLAGPLHGLANOEVLMLOLOKAXXAGADAS 300  
 QY 328 LRDIYINNTLNSGRVPGYGAHVLAKTDPRTYTCOREFALKHLPNDEPKELVLAOLYKTIYPNV 387  
 DB 301 LRDIYINNTLNSGRVPGYGAHVLAKTDPRTYTCOREFALKHLPGDPMKELVLAOLYKTIYPNV 360  
 QY 388 LLEOGAKAPNPVNDVHSGVLLQYGGTEKNYTYVLEGVSRALGVLAQLIWSRALGFPLE 447  
 DB 361 LLEOGAKAPNPVNDVHSGVLLQYGGTEKNYTYVLEGVSRALGVLAQLIWSRALGFPLE 420  
 QY 448 RPKSMSTEGIM 458  
 DB 421 RPKSMSTEGIM 431

RESULT 4  
 CISTY\_CABEL STANDARD; PRT; 468 AA.  
 AC P34575;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Probable citrate synthase, mitochondrial precursor (EC 2.3.3.1).

GN T20G5.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderiinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Berks M., Smith A.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate = citrate +  
 COA.  
 CC -1- PATHWAY: Tricarboxylic acid cycle.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).  
 CC -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS  
 CC CAPABLE OF OXIDATIVE METABOLISM.  
 CC -1- SIMILARITY: Belongs to the citrate synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z30423; CAA83004.1; -  
 DR PIR: S42370; S42370.  
 DR HSSP: P23007; 2CSC.  
 DR WormRep: T20G5.2; CE00513.  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR Pfam: PF00285; Citrate\_synth; 1.  
 DR PRINTS: PR00143; CITRATESYNTHASE.  
 DR PROSITE: PS00480; CITRATE SYNTHASE; 1.  
 KM Hypothetical protein; Transferase; Tricarboxylic acid cycle;  
 KM Mitochondrion; Transit peptide.  
 FT TRANSIT 1 468  
 FT CHAIN 1 468  
 FT ACT\_SITE 303 303  
 FT ACT\_SITE 349 349  
 FT ACT\_SITE 404 404  
 SQ SEQUENCE 468 AA: 51540 MW: CDE40CC45729B9BEA CRC64;

Query Match 69.5%; Score 1693; DB 1; Length 468;  
 Best Local Similarity 68.3%; Pred. No. 8.3e-120;  
 Matches 315; Conservative 62; Mismatches 82; Indels 2; Gaps 1;

QY 1 MALLTAAARLGRKNA--SCLVLAARHNASSTNLKDIADLIPKQARIKTFROOHGKT 58  
 DB 1 MSLSGMAIRRLITKGYIPVCQVAPLSTABGSTNLKELVSKLIPAHNAKYSRTTERGST 60  
 QY 59 VGOQITVDMYGMGMKGLVETSYLDPDEGIRRGFSIPKQKLLPRKAGGEPLEPG 118  
 DB 61 VGOQITVDMYGMGMKGLVETSYLDPDEGIRRGFSIPKQKLLPRKAGGEPLEPG 120  
 QY 119 LFWLVTGHIPTPEOYSMLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAVYALNSE 178  
 DB 121 LFWLVTGHIPTPEOYSMLSKEMAKRAALPSHVVTMLDNFPTNLHPMSQLSAVYALNSE 180  
 QY 179 SNFARAAYAGISRTKYMWELIVEDSMOLAIPCVAAKIYNTLIRBEGSIGAIDSNLDMSH 238  
 DB 181 SNFARAAYAGISRTKYMWELIVEDSMOLAIPCVAAKIYNTLIRBEGSIGAIDSNLDMSH 240  
 QY 239 NFNMLGYTDHOFTELRLYLTIHSHDEGNSAHTSHLVGSALSDPYLFAAAMNGLAG 298  
 DB 241 NFNMLGYTDHOFTELRLYLTIHSHDEGNSAHTSHLVGSALSDPYLFAAAMNGLAG 300  
 QY 299 PLHGLANOEVLMLOLOKEVGKDYSDKLRDIYINNTLNSGRVPGYGAHVLAKTDPRTY 358  
 DB 301 PLHGLANOEVLMLOLOKEVGKDYSDKLRDIYINNTLNSGRVPGYGAHVLAKTDPRTY 360  
 QY 359 COREFALKHLPNDEPKELVLAOLYKTIYPNVLEOGAKAPNPVNDVHSGVLLQYGGTEKN 418

```

Db      361 CREFALKHLPNDJFKLVSTLYKTPGILLFQGRKAKNPWPNVDSHSGVLLQYFGWTEKS 420
OY      419 YTVTFGVSRALGVAQLIWSRALGFPLEPRKSMSTEGLMK 459
Db      421 FTVTFGVSRALGCLSLIWMARGMGLPERKSHSTDGLIK 461

RESULT 5
CISY_EMENT
ID      CISY_EMENT      STANDARD:      PRT:      474 AA.
AC      000098;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Citrate synthase, mitochondrial precursor (EC 2.3.3.1).
GN      CITR.
OS      Emericella nidulans (Aspergillus nidulans).
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OX      Eurotiiales; Trichocomaceae; Emericella.
RN      NCBI_TaxID=162425;
RP      SEQUENCE FROM N.A.
RC      STRAIN=FGSC 4;
RX      MEDLINE=97306446; PubMed=9163747;
RA      Park B.W., Han K.H., Lee C.Y., Lee C.H., Maeng P.J.;
RT      "Cloning and characterization of the citA gene encoding the
RT      mitochondrial citrate synthase of Aspergillus nidulans."
RL      Mol. Cells 7:290-295(1997).
CC      -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate -> citrate +
CC      CoA.
CC      -1- PATHWAY: Tricarboxylic acid cycle.
CC      -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC      -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS
CC      CAPABLE OF OXIDATIVE METABOLISM.
CC      -1- SIMILARITY: Belongs to the citrate synthase family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: U89675; AAC49728.1; -
DR      HSSP: P23007; 2CSC.
DR      InterPro: IPR002020; Citrate_synth.
DR      Pfam: PF00285; citrate_synth; 1.
DR      PRINTS: PR00143; CITRSMTHASE.
DR      PROSITE: PS00480; CITRATE_SYNTHASE; 1.
KM      Transferase: Tricarboxylic acid cycle; Mitochondrion; Transit peptide.
FT      TRANSIT      1      35      MITOCHONDRION (POTENTIAL).
FT      CHAIN      36      474      CITRATE SYNTHASE.
FT      ACT_SITE      310      310      BY SIMILARITY.
FT      ACT_SITE      356      356      BY SIMILARITY.
FT      ACT_SITE      411      411      BY SIMILARITY.
SO      SEQUENCE      474 AA; 52223 MW; E2E86892ACB5398B CRC64;

Query Match      63.5%; Score 1547.5; DB 1; Length 474;
Best Local Similarity 62.2%; Pred. No. 7.3e-109;
Matches 294; Conservative 68; Mismatches 96; Indels 15; Gaps 3;

OY      6 AAARLGTAKNSCLVLAAR-----HASASSTNLKDIADLIPKQAKIKTR 52
Db      2 ASTLRSLSTALRSSTLAKPVQSAFNGLCRYSTGKTKSLKTEPADKLPGLELVKLR 61
OY      53 QOHGKTVVQGLTVDMYGGMRGAKGLVETSVLADDEGIRFGFSIPECOKLIPKAKGE 112
Db      62 KKHGKRVIGELTLDQAYGARGVKCLVWGSLDSEEGIRFGGLTIPKQKLLPRAPGE 121
OY      113 EPLRGLFWLVTGHIPTBEQVSWLSKEMAKRALPLSHVYVLDNFPVTLHPMSQLSAV 172

```

```

Db      122 EPLRGLFWLTLTGEVSEQYRDLSEAWAARSDDLPRFIEELIDRVPSTLHPMAQFSLAV 181
OY      173 TALNSESNNRARAAGCISRTKYMELIEDSMILAKLPQVAKIYRNLYREGSGIGALDS 232
Db      182 TALEHESAPAKATAPKINKKDWNTWTFEDSMILAKLPITIAKIRNNFKDK -VAPLOK 240
OY      233 NIDMSHNFNMGLYTDHQ-FTELRLYLTLTHSDHEGANSATSHLVGSALSAPYLSFNA 291
Db      241 DKDYSNLANOGLGFADNKNFVELMLRYLTLTHSDHEGANSATTHLVGSALSAPLSLA 300
OY      292 ANGLAGPLHLGLANDVWLTLQLOKENGKDVSDKLRDYIWNPLNSGRVDPGYHAVALR 351
Db      301 GLNGLAGPLHLGLANDVWLTLQLOKENGKDVSDKLRDYIWNPLNSGRVDPGYHAVALR 360
OY      352 KNDPRYTQOREALHLPNDPMPKFLVAOLYKRVPNVLLFQGRKAKNPWPNVDSHSGVLLQY 411
Db      361 KNDPRYTQOREALHLPNDPMPKFLVAOLYKRVPNVLLFQGRKAKNPWPNVDSHSGVLLQY 420
OY      412 YGTEMNMYTVTFGVSRALGVAQLIWSRALGFPLEPRKSMSTEGLMKFPVDSK 464
Db      421 YGLTERNYTVTFGVSRALGVAQLIWSRALGFPLEPRKSMSTEGLMKFPVDSK 473

RESULT 6
CISY_ASPIG
ID      CISY_ASPIG      STANDARD:      PRT:      475 AA.
AC      P51044;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Citrate synthase, mitochondrial precursor (EC 2.3.3.1).
GN      CIT-1.
OS      Aspergillus niger.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OX      Eurotiiales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus.
RN      NCBI_TaxID=5061;
RP      SEQUENCE FROM N.A.
RC      STRAIN=WMU-2223L;
RA      Oshida Y., Miyake K., Kanayama S., Kitamura K., Usami S.;
RL      Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
CC      -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate -> citrate +
CC      CoA.
CC      -1- PATHWAY: Tricarboxylic acid cycle.
CC      -1- SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).
CC      -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS
CC      CAPABLE OF OXIDATIVE METABOLISM.
CC      -1- SIMILARITY: Belongs to the citrate synthase family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: D63376; BAA09691.1; -
DR      HSSP: P23007; 2CSC.
DR      InterPro: IPR002020; Citrate_synth.
DR      Pfam: PF00285; citrate_synth; 1.
DR      PRINTS: PR00143; CITRSMTHASE.
DR      PROSITE: PS00480; CITRATE_SYNTHASE; 1.
KM      Transferase: Tricarboxylic acid cycle; Mitochondrion; Transit peptide.
FT      TRANSIT      1      475      CITRATE SYNTHASE.
FT      CHAIN      310      310      BY SIMILARITY.
FT      ACT_SITE      356      356      BY SIMILARITY.
FT      ACT_SITE      411      411      BY SIMILARITY.
SO      SEQUENCE      475 AA; 52153 MW; F93525B3F1FCB3F CRC64;

Query Match      62.7%; Score 1526.5; DB 1; Length 475;
Best Local Similarity 62.5%; Pred. No. 2.8e-107;

```





DR pfam: PR00285; citrate\_synth. 1.  
 DR PRINTS; PR00143; CITRATESYNTHASE.  
 DR PROSITE; PS00480; CITRATE\_SYNTHASE; 1.  
 KW transferase; Tricarboxylic acid cycle; Mitochondrion; Transit peptide;  
 Multigene family.  
 FT TRANSIT 1 33 MITOCHONDRION (POTENTIAL).  
 FT CHAIN 34 469 CITRATE SYNTHASE.  
 FT ACT\_SITE 352 352 BY SIMILARITY.  
 FT CONFLICT 351 351 G -> A (IN REF. 1).  
 SO SEQUENCE 469 AA; 51988 MW; 14AFDDE07D345FD CRC64;  
 Query Match 61.6%; Score 1501.5; DB 1; Length 469;  
 Best Local Similarity 62.2%; Pred. No. 2.1e-105;  
 Matches 283; Conservative 79; Mismatches 88; Indels 5; Gaps 3;  
 QY 2 ALTAARLIGTKNASCVLAAARHASSSTNLKDLADLPKROAKIKTFROQHGKTVG 61  
 DB 10 ALRRSILHLSRQTA---FTAAACYSKQTLKERAPELLPENIEIKALKRHKSKYVD 66  
 QY 62 QITVDMMGKMGKGLVETSYLDPDEGIRFRGSEIPRQKILPRAKGGEPLPGLFW 121  
 DB 67 KYLDQVYGARGIKCLVWEGSVLDAEEGIRFRGKTIPEQELLPRAPGKPEPLPGLFW 126  
 QY 122 LVTGHIPIEEQVSWLSEKEMARALPSHYVTMLDNFPTNLHPMSQLSAVTAALNSESNF 181  
 DB 127 LLTGEVPSBOQVRLSAEMARSDVPKTEELIDPCPSDLHPMAQLSTAVTALERTSSP 186  
 QY 182 ARAYAGISRTKYMELIYEDSMDLAKPCVAAKIYRNLREGSGIGALDSINLDMSHNFT 241  
 DB 187 ARYAGINKKEKWEVTFEDSMDLAKPLTARIQVNFK-GGKAAVQKMDYFNF 245  
 QY 242 NMIGYDHO-FTELTLYLTLSHDEGCVNSANTSHLVSAISDPLSFRAAMNGIAGPL 300  
 DB 246 NDLGFDNDVEFELLTYLTHDHEGCVNSANTSHLVSAISDPLSFRAAMNGIAGPL 305  
 QY 301 HGLANQEVLMLOLOKEVGDVDEKLDYINWNTLNSGVVGYGHAUVRKTDPRYTCO 360  
 DB 306 HGLANQEVLMLOLOKEVGDVDEKLDYINWNTLNSGVVGYGHAUVRKTDPRYTCO 365  
 QY 361 REFALKHLPNDPMFKLQALYKIVPVLLLEOGKAKNPVNDASHGVLLQYGMTEMYNY 420  
 DB 366 RKFAQHLPEDPMFKLQALYKIVPVLLLEOGKAKNPVNDASHGVLLQYGMTEMYNY 425  
 QY 421 TVLFGSRALGVLAOLINSRAGLPLEPKSKMSTE 455  
 DB 426 TVLFGSRALGVLAOLINSRAGLPLEPKSKMSTE 460  
 RESULT 10  
 CITSY\_YEAST STANDARD; PRT; 479 AA.  
 AC P00890;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE citrate synthase, mitochondrial precursor (EC 2.3.3.1).  
 GN CIT1 OR LIS5 OR GLU3 OR YNR001C OR N2019.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 NX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=D273-10B;  
 RC Linder P., Plueckthun A.;  
 RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D273-10B;  
 RC MEDLINE=85003587; PubMed=6090126;  
 RX Luisa M., Suda K., Schatz G.;  
 RT "Isolation of the nuclear yeast genes for citrate synthase and  
 fifteen other mitochondrial proteins by a new screening method.";

RL EMBO J. 3:1773-1781(1984).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / FY1679;  
 RX MEDLINE=95208356; PubMed=7900425;  
 RA Verhasselt P., Aert R., Voet M., Volckaert G.;  
 RT "Twelve open reading frames revealed in the 23.6 kb segment flanking  
 the centromere on the Saccharomyces cerevisiae chromosome XIV right  
 arm.";  
 RT Yeast 10:1355-1361(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GRF88;  
 RX MEDLINE=95028151; PubMed=7941739;  
 RA Lalo D., Stettler S., Mariotte S., Gendreau E., Thuriaux P.;  
 RT "Organization of the centromeric region of chromosome XIV in  
 Saccharomyces cerevisiae.";  
 RT Yeast 10:523-533(1994).  
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate = citrate +  
 COA.  
 CC -1- PATHWAY: Tricarboxylic acid cycle.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS  
 CAPABLE OF OXIDATIVE METABOLISM.  
 CC -1- SIMILARITY: Belongs to the citrate synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC EMBL: 223259; CAA80781.1; -  
 DR EMBL: X00782; CAA25359.1; -  
 DR EMBL: X77395; CAA54569.1; -  
 DR EMBL: Z71616; CAA96277.1; -  
 DR PIR: S35390; YKBY.  
 DR HSSP: P23007; 2CSC.  
 DR SGD: S0005284; CIT1.  
 DR GO: GO:0005739; C:mitochondrion; IDA.  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR pfam: PR00285; citrate\_synth.  
 DR PRINTS; PR00143; CITRATESYNTHASE.  
 DR PROSITE; PS00480; CITRATE\_SYNTHASE; 1.  
 DR transferase; Tricarboxylic acid cycle; Mitochondrion; Transit peptide;  
 KW Multigene family.  
 KM TRANSIT 1 37 MITOCHONDRION.  
 FT CHAIN 38 479 CITRATE SYNTHASE.  
 FT ACT\_SITE 312 312 BY SIMILARITY.  
 FT ACT\_SITE 358 358 BY SIMILARITY.  
 FT ACT\_SITE 413 413 BY SIMILARITY.  
 FT CONFLICT 58 58 E -> Q (IN REF. 2).  
 FT CONFLICT 78 78 E -> E (IN REF. 2).  
 SO SEQUENCE 479 AA; 53360 MW; 280661B1CB248F14 CRC64;  
 Query Match 61.6%; Score 1499.5; DB 1; Length 479;  
 Best Local Similarity 62.2%; Pred. No. 3e-105;  
 Matches 280; Conservative 76; Mismatches 89; Indels 5; Gaps 3;  
 QY 20 VLAARH-ASASSSTNLKDLADLPKROAKIKTFROQHGKTVGQITVDMMGKMGKGL 78  
 DB 31 LNAHRYSSASEQTLEKFAEITIPAKAEIKFKKKGKTVIGEVLLLEQAYGMRIGKL 90  
 QY 79 VVETSVLDPDEGIRFRGSEIPRQKILPRAKGGEPLPGLFWLVTGHIPIEEQVSWLS 138  
 DB 91 VMEGSLVDPDEGIRFRGKRTIPERQRELPRKAEGSTPEPLPALFWLLTGEIPIIDAQYKALS 150  
 QY 139 KEMAKRAALPSHYVTMLDNFPTNLHPMSQLSAVTAALNSESNFARAYAGISRTKYMELI 198  
 DB 151 ADLAARSEIPEHVITQLDLSLPLDHPMAQLSTAVTALSESSEKFAKAYAGVSKKEYWSTY 210



QY 199 YEDSMDLIAPCAVAAKIYRNLYREGSIGAIDSNLMDNSHNTNMLGYTDHOFTELRLY 258  
D 211 FEDSLDILGKLPVIAASKIRYRVFKDK - ITSTDPMADYCKMLAQLGVENNDFIDMLRLY 269  
QY 259 LTHSDHEGCVNSAHTSHLSALSADPLSFPAAMNGLAGPLHGLANOVLWLTOLKE 318  
D 270 LTHSDHEGCVNSAHTSHLSALSADPLSFPAAMNGLAGPLHGLANOVLWLTOLKE 329  
QY 319 VGRVSDKRLDYLWNTLNSGRVPGYGAVALRKTDPRTQOREFALKHNDMPKLYA 378  
D 330 VKGYSKETIKRYMDLINAARVPGYGAVALRKTDPRTQOREFALKHNDMPKLYA 389  
QY 379 QLYKIVPNVLEQKAKNPNDVAHSGVLLQYGMENYTYVLEFVSRAVLAQLIM 438  
D 390 TIVAVPGLVFKHCKTKNPNVDSHSGVLLQYGLTFASFTYVLFVARRAIGVLPOLI 449  
QY 439 SRAIGFLEPKSMSTE---GLAKFVDSKS 465  
D 450 DRAVGAPLERPKSFSTEKRYKELVKIESKN 479

RESULT 11

CISY\_ARATH STANDARD; PRT: 473 AA.  
AC P20115; Q64869; Q8LE36; Q94EY6;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Citrate synthase, mitochondrial precursor (EC 2.3.3.1).  
GN ATG44350 OR F411.16.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91370823; PubMed=2491664;  
RA Unger E.A., Hand J.M., Cashmore A.R., Vasconcelos A.C.;  
RT "Isolation of a cDNA encoding mitochondrial citrate synthase from  
RT Arabidopsis thaliana."  
RL Plant Mol. Biol. 13:411-418(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shee T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,  
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,  
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Cressy T.H.,  
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
RA Niernman W.C., White O., Eissen J.A., Salzberg S.L., Fraser C.M.,  
RA Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
RT thaliana."  
RL Nature 402:761-768(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
RT "RIKEN Arabidopsis full length cDNA clones (RAPFs) sequenced by the  
RT SSP consortium (Salk/Stanford/GCFC)."  
RL Submitted (May-2001) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.;  
RT "Full-length cDNA from Arabidopsis thaliana."  
RL Submitted (Mar-2002) to the EMBL/Genbank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate -> citrate +  
CC CoA.  
CC -1- PATHWAY: Tricarboxylic acid cycle.

CC -1- SUBUNIT: Homodimer.  
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
CC -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS  
CC CAPABLE OF OXIDATIVE METABOLISM.  
CC -1- SIMILARITY: Belongs to the citrate synthase family.  
CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts  
CC in positions 97; 105; 119; 125; 137; 149; 263; 271; 386; 390; 393;  
CC 398 and 403.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; X17528; CAA35570.1; ALT\_FRAME.  
CC EMBL; AC004521; AAC16084.2; -  
CC EMBL; AF387018; AAK62463.1; -  
CC EMBL; AY085647; AAM62868.1; -  
CC HSPSP; P23007; 2CSC.  
CC InterPro: IPR002020; Citrate-synt.  
CC Pfam: PF00285; citrate-synt. 1.  
CC PRINTS: PR00143; CITRATESYNTHASE.  
CC PROSITE: PS00480; CITRATE SYNTHASE; 1.  
CC Transferrase: Tricarboxylic acid cycle; Mitochondrion; Transit peptide;  
CC Multigene family.  
CC KM TRANSIT  
CC FT CHAIN 1 16  
FT ACT\_SITE 307 473 MITOCHONDRION (POTENTIAL).  
FT ACT\_SITE 307 307 CITRATE SYNTHASE.  
FT ACT\_SITE 353 353 BY SIMILARITY.  
FT ACT\_SITE 408 408 BY SIMILARITY.  
FT CONFLICT 17 17 V -> VO (IN REF. 1).  
FT CONFLICT 24 24 S -> N (IN REF. 4).  
FT CONFLICT 51 52 QQ -> HK (IN REF. 1).  
FT CONFLICT 111 111 V -> AL (IN REF. 1).  
FT CONFLICT 155 155 P -> S (IN REF. 1).  
FT CONFLICT 256 257 KV -> R (IN REF. 1).  
FT CONFLICT 374 374 L -> H (IN REF. 1).  
FT CONFLICT 382 384 LVS -> CC (IN REF. 1).  
FT CONFLICT 447 449 ALG -> ELI (IN REF. 1).  
SO SEQUENCE 473 AA; 52654 MM; DF747D3DA8943FFB CRC64;  
  
Query Match 61.5%; Score 1498.5; DB 1; Length 473;  
Best Local Similarity 62.6%; Pred. No. 3.5e-105;  
Matches 270; Conservative 77; Mismatches 83; Indels 1; Gaps 1;  
  
QY 27 SASSTNLKDLADLPKEOARIKTFROOQKTVVQOIVDMYGMGRMGKGLVETSYLD 86  
D 34 SSTDLKLSQGLPEODRLKRLKSHKGVOLGNIVDMVIGMRMGTLGMLTSLD 93  
QY 87 PDEGIRFPGSIPKQKLPKAKGGEPLPGLTFLVLTGHIPIEBEVSMLSEKAKAA 146  
D 94 PEEGIRFPGSIPKQKLPKAKGGEPLPGLTFLVLTGHIPIEBEVSMLSEKAKAA 153  
QY 147 LPSHVYMLDNFPPNLPMSOASAATVLSNESNFRARYAGISQITWELIYEDSM 206  
D 154 VPDVYVNAIDALPSTAHPTQFASGVMALOYOSEFOKAYENGIRHKSWEPTTECC 213  
QY 207 AKLPCVAAKIYRNLYREGSIGAIDSNLMDNSHNTNMLGYTDHOFTELRLYLTISDHE 266  
D 214 ARVPVAAVYRMYKNGDSIPS -DKSLDYGANSNMGFDEKVELMRLYITISDHE 272  
QY 267 GGNVSAHTSHLVGSLSPYTSFAAAMNGLAGPLHGLANOVLWLTOLKEVKDVSDE 326  
D 273 GGNVSAHTSHLVGSLSPYTSFAAAMNGLAGPLHGLANOVLWLTOLKEVKDVSDE 332  
QY 327 KLRBYVNTLNSGRVPGYGAVALRKTDPRTQOREFALKHNDMPKLYAQLIYV 386  
D 333 QLKEIVYVNTLNSGRVPGYGAVALRKTDPRTQOREFALKHNDMPKLYAQLIYV 392  
QY 387 VLEEGKAKNPVNDVAHSGVLLQYGMENYTYVLEFVSRAVLAQLIWSRALGFP 446



QY 5 TAAATLGC-----TKNASCLVLAARHASASTNLKDLIADLPKEQARIKTFEQQHGTIV 60  
 DB 8 TALSLRSRAGQOSLSNSVRLQMOSSADLDHLSQLEKMEPEQERLKKVSDLGKQOL 67  
 QY 61 GQITVDMYGMGRMGKGLVETSVLDPDEGIRFEGFSIPECOKLLPKAKGEEPLPEGLF 120  
 DB 68 GNITIDVYGMGRMGKGLVETSVLDPDEGIRFEGFSIPECOKLLPKAKGEEPLPEGLF 127  
 QY 121 WLVYGHIPTEBOVSMLSKEAKRRAALPSHYVTMLDNPTNLHPMSQLSAAVTALNSSN 180  
 DB 128 WLLLGKVPSEQYDGLSKELDRATVPDYKADALPVSHPMTQFASGVMALQVQSE 187  
 QY 181 FARVAQGISRTKYMELIYEDSDMDLIARLPCVAKIYINLYREGSGIAIDSNLDMSENF 240  
 DB 188 FOENAEKGIHKSKEPSEDSLSNLIAVPVAAVYQRIKDKIIRK-DSDLDYGNF 246  
 QY 241 TNNMGTYDHOFTELTRLYLTTHSDHEGCVNSAHTSHLVGALSALDPYLSFAAAMNLAGPL 300  
 DB 247 SHMLGFDDPKMLRLMRLVYTHSHDEGCVNSAHTSHLVGALSALDPYLSFAAAMNLAGPL 306  
 QY 301 HGLANQEVLYMVLTOLOKREKVDSEKLDYITMNTLNSGRVPGYGHAVLRKTDPRYTCQ 360  
 DB 307 HGLANQEVLYMVLTOLOKREKVDSEKLDYITMNTLNSGRVPGYGHAVLRKTDPRYTCQ 366  
 QY 361 REFALKHLPPNPMFELVAQLYKIVPNVLEOGKAKNPMPNDASHGVLLQYGYGTEMNY 420  
 DB 367 REFALKHLPPNPMFELVAQLYKIVPNVLEOGKAKNPMPNDASHGVLLQYGYGTEMNY 426  
 QY 421 TLFEGVSRAVLQALQIWSRALGFPLEPRKSMTEGLMK 459  
 DB 427 TLFEGVSRAVLQALQIWSRALGFPLEPRKSMTEGLMK 465

## RESULT 14

CIST\_YEAST STANDARD: PRT: 460 AA.  
 ID CIST\_YEAST AC P08679;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Citrate synthase, peroxisomal (EC 2.3.3.1).  
 GN C172 OR YCR005C OR YCR5C OR YCR043.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 CX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92254505; PubMed=1580102;  
 RA Bileau N., Fremaux C., Hebrard S., Menara A., Algie M., Crouzet M.;  
 RT "The complete sequence of a 10.8kb fragment to the right of the  
 RT chromosome III centromere of Saccharomyces cerevisiae.";  
 RL Yeast 8:61-70(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87089811; PubMed=3540614;  
 RA Rosenkrantz M., Alam T., Kim K.-S., Clark B.J., Stere P.A.,  
 RA Cuarente L.P.;  
 RT "Mitochondrial and nonmitochondrial citrate synthases in  
 RT Saccharomyces cerevisiae are encoded by distinct homologous genes.";  
 RL Mol. Cell. Biol. 6:4509-4515(1986).  
 RN [3]  
 RP SEQUENCE OF 1-24 FROM N.A.  
 RX MEDLINE=91094853; PubMed=1986232;  
 RA Liao X., Small W.C., Stere P.A., Butow R.A.;  
 RT "Intramitochondrial functions regulate nonmitochondrial citrate  
 RT synthase (CIT2) expression in Saccharomyces cerevisiae.";  
 RL Mol. Cell. Biol. 11:38-46(1991).  
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate -> citrate +  
 CC COA.  
 CC -1- PATHWAY: Tricarboxylic acid cycle.  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.

CC -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS  
 CC CAPABLE OF OXIDATIVE METABOLISM.  
 CC -1- SIMILARITY: Belongs to the citrate synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: Z11113; CA77442.1;  
 DR EMBL: M14686; AAA34497.1;  
 DR EMBL: M54982; AAA34498.1;  
 DR EMBL: X59720; CA42342.1;  
 DR PIR: A25393; YKBYC.  
 DR HSSP: P23007; 2CSC.  
 DR SGD: S0000598; CIT2.  
 DR GO: GO:0005777; C:peroxisome; IDA.  
 DR InterPro: IPR002020; Citrate\_synth.  
 DR Pfam: PF00285; citrate\_synth.  
 DR PRINTS: PR00143; CITRATESINASE.  
 DR PROSITE: PS00480; CITRATE\_SYNTHASE; 1.  
 KM Transferrase; Tricarboxylic acid cycle; peroxisome; Multigene family.  
 FT ACT\_SITE 293 293 BY SIMILARITY.  
 FT ACT\_SITE 339 339 BY SIMILARITY.  
 FT ACT\_SITE 394 394 BY SIMILARITY.  
 FT SITE 458 460 MICROBODY TARGETING SIGNAL (POTENTIAL).  
 SQ SEQUENCE 460 AA: 51413 MW; AB2F6AD9A9399EF CRC64.

Query Match 58.5%; Score 1425; DB 1; Length 460;  
 Best local similarity 59.9%; Pred. No. 1.1e-99;  
 Matches 264; Conservative 72; Mismatches 101; Indels 4; Gaps 2;

QY 27 SASSTNLKDLIADLPKEQARIKTFEQQHGTIVGQITVDMYGMGRMGKGLVETSVLD 86  
 DB 20 SSOEKLTKERSEIPIHAQDVQFVKHKGTKISDVLEQYVGGMRGIPSGVMEGSYLD 79  
 QY 87 PDEGIRFEGFSIPECOKLLPKAKGEEPLPEGLFVLYTGHIPTEBOVSMLSKEAKRAA 146  
 DB 80 PEDGIRFEGFSIPECOKLLPKAKGEEPLPEGLFVLYTGHIPTEBOVSMLSKEAKRAA 139  
 QY 147 LPSHYVTMLDNPTNLHPMSQLSAAVTALNSESNFARVAQGISRTKYMELIYEDSDMDLI 206  
 DB 140 LPSHYVTMLDNPTNLHPMSQLSAAVTALNSESNFARVAQGISRTKYMELIYEDSDMDLI 199  
 QY 207 AKLPCVAAKIYINLYREGSGIGAIDSNDLWSHNTMLGYTDHOFTELTRLYLTTHSDHE 266  
 DB 200 GKLPIVIAKIYINLYREGSGIGAIDSNDLWSHNTMLGYTDHOFTELTRLYLTTHSDHE 258  
 QY 267 GGNVSAHTSHLVGALSALDPYLSFAAAMNLAGPLHGLANQEVLYMVLTOLOKREKVDSE 326  
 DB 259 GGNVSAHTSHLVGALSALDPYLSFAAAMNLAGPLHGLANQEVLYMVLTOLOKREKVDSE 318  
 QY 327 KLRDYITMNTLNSGRVPGYGHAVLRKTDPRYTCQREFALKHLPPNPMFELVAQLYKIVPN 386  
 DB 319 TLEKTLMTLNSGRVPGYGHAVLRKTDPRYTCQREFALKHLPPNPMFELVAQLYKIVPN 378  
 QY 387 VLEGGKAKNPMPNDASHGVLLQYGYGTEMNYTVLFEGVSRAVLQALQIWSRALGFP 446  
 DB 379 VLEGGKAKNPMPNDASHGVLLQYGYGTEMNYTVLFEGVSRAVLQALQIWSRALGFP 438  
 QY 447 ERPKSMSTE---GLMKFPVDSK 464  
 DB 439 ERPKSYSTEKYEKLVKNIESK 459

## RESULT 15

CIST\_FRAAN STANDARD: PRT: 469 AA.  
 ID CIST\_FRAAN AC P8372;  
 DT 28-FEB-2003 (Rel. 41, Created)

Search completed: September 5, 2003, 20:08:56  
Job time : 25 secs

DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Citrate synthase, mitochondrial precursor (EC 2.3.3.1).  
 GN MGSI.  
 OS *Fragaria ananassa* (Strawberry).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;  
 OC eustosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
 OX NCBI\_TaxID=3747;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-46.  
 RC STRAIN=CV. ELISANTA; TISSUE=FRUIT;  
 RA Iannetta R.P., Davies H.V., Medina Escobar N., Ross H., Souleyre E.,  
 RA Hancock R.D., Davies H.V.;  
 RT "Mitochondrial citrate synthase and -malate dehydrogenase in ripening  
 RT strawberry fruit".  
 RL Physiol. Plantarum 0:0-0(2002).  
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate = citrate +  
 CC CoA.  
 CC -1- PATHWAY: Tricarboxylic acid cycle.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- MISCELLANEOUS: Citrate synthase is found in nearly all cells  
 CC capable of oxidative metabolism.  
 CC -1- SIMILARITY: Belongs to the citrate synthase family.  
 DR InterPro:IPR002020, Citrate\_synth.  
 DR Pfam:PF00285, Citrate\_synth.1.  
 DR PROSITE:PS00480, CITRATE\_SYNTHASE.1.  
 FT TRANSIT: Tricarboxylic acid cycle; Mitochondrion; Transit peptide.  
 FT TRANSIT 1 28 MITOCHONDRION.  
 FT CHAIN 29 469 CITRATE SYNTHASE.  
 FT ACT\_SITE 304 304 BY SIMILARITY.  
 FT ACT\_SITE 350 350 BY SIMILARITY.  
 FT ACT\_SITE 405 405 BY SIMILARITY.  
 SEQUENCE 469 AA; 52282 MW; CD80AEDD06246F65 CRC64;

Db 1 MALLTAARLFGAKNASCLVLAARHASASSTNLKDI LADLIPKEQARIKTFROOHGNTVV 60

**THIS PAGE BLANK (USPTO)**

OY 61 GQITVDAMYGMRGKGLVETSVLDPDEGIRFRGFSIPBECOKLPRKAGGEEPLPEGLF 120  
 DB 61 GQITVDAMYGMRGKGLVETSVLDPDEGIRFRGFSIPBECOKLPRKAGGEEPLPEGLF 120  
 OY 121 WLVLTGHIPTBEOVSWLSKEMAKRAALPSHVYTMNDNPTNLHPMSQLSAAVTALNSNSN 180  
 DB 121 WLVLTGHIPTBEOVSWLSKEMAKRAALPSHVYTMNDNPTNLHPMSQLSAAVTALNSNSN 180  
 OY 181 FARATAOGISRTKYVELLYEDSMOILAPCAAKIYRNLYREGSGTALDINSNDMSNF 240  
 DB 181 FARATAOGISRTKYVELLYEDSMOILAPCAAKIYRNLYREGSGTALDINSNDMSNF 240  
 OY 241 TMLAGYTHOFTLRLYLTITSDHEGNGVSAHTSHLVGALSADPYLSFAAAMNGLAGPL 300  
 DB 241 TMLAGYTHOFTLRLYLTITSDHEGNGVSAHTSHLVGALSADPYLSFAAAMNGLAGPL 300  
 OY 301 HGLANOEVLWLTOLQKEVGDVSDKLDYIMNTLNSGRVVPYGAHVLKRTDPRYTCQ 360  
 DB 301 HGLANOEVLWLTOLQKEVGDVSDKLDYIMNTLNSGRVVPYGAHVLKRTDPRYTCQ 360  
 OY 361 REFALKHLPNDEPFVLVQLYKIVPNVLLGOGAKNPNPNVDAHSGVLLQYGMTEMY 420  
 DB 361 REFALKHLPNDEPFVLVQLYKIVPNVLLGOGAKNPNPNVDAHSGVLLQYGMTEMY 420  
 OY 421 TVLFGVSRLGVLAOLIMSRALGFPLERPKSMSTEGLMKFVDSK 464  
 DB 421 TVLFGVSRLGVLAOLIMSRALGFPLERPKSMSTEGLMKFVDSK 464

## RESULT 2

542370  
 C:Date: 07-Oct-1994 #sequence\_revision 10-Nov-1995 #extl\_change 05-May-2000  
 C:Accession: S42370  
 R:Smith, A.  
 submitted to the EMBL Data Library, March 1994  
 A:Reference number: S42368  
 A:Accession: S42370  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-468 <SMI>  
 A:Cross-references: EMBL:230423; NID:9458479; PID:9458482  
 C:Genetics:  
 A:Introns: 69/3; 202/3; 309/3  
 C:Superfamily: citrate (si)-synthase  
 C:Keywords: carbon-carbon lyase; coenzyme A; mitochondrion; oxo-acid-lyase

Query Match 69.5%; Score 1693; DB 2; Length 468;

Best Local Similarity 68.3%; Pred. No. 3.9e-120; Mismatches 82; Indels 2; Gaps 1;

OY 1 MALLTAARLLGTRKN--SCLVLAARHASSSTNLDIADLIPKBOARKITRROOCHT 58  
 DB 1 MSLSGMARLRLITKGVIPCOVAPLSTSAEGSTNLEKVLSPKIPAHNAKKSFRTEHGS 60  
 OY 59 VVGQITVDAMYGMRGKGLVETSVLDPDEGIRFRGFSIPBECOKLPRKAGGEEPLPEG 118  
 DB 61 VVGQITVDAMYGMRGKGLVETSVLDPDEGIRFRGFSIPBECOKLPRKAGGEEPLPEG 120  
 OY 119 LFWLLVTGHIPTBEOVSWLSKEMAKRAALPSHVYTMNDNPTNLHPMSQLSAAVTALNS 178  
 DB 121 LFWLLVTGHIPTBEOVSWLSKEMAKRAALPSHVYTMNDNPTNLHPMSQLSAAVTALNS 180  
 OY 179 SNFARAYAGISRTKYVELLYEDSMOILAPCAAKIYRNLYREGSGTALDINSNDMSH 238  
 DB 181 SKFAGAAARVAKASWEYALIEDSMOILAPCAAKIYRNLYREGSGTALDINSNDMSH 240  
 OY 239 NFWTMLGTYDHOFTLRLYLTITSDHEGNGVSAHTSHLVGALSADPYLSFAAAMNGLAG 298  
 DB 241 NFWTMLGTYDHOFTLRLYLTITSDHEGNGVSAHTSHLVGALSADPYLSFAAAMNGLAG 300  
 299 PLHGLANOEVLWLTOLQKEVGDVSDKLDYIMNTLNSGRVVPYGAHVLKRTDPRYT 358

DB 301 PLHGLANOEVLWLTOLQKEVGDVSDKLDYIMNTLNSGRVVPYGAHVLKRTDPRYT 360  
 OY 359 COREFALKHLPNDEPFVLVQLYKIVPNVLLGOGAKNPNPNVDAHSGVLLQYGMTEMY 418  
 DB 361 COREFALKHLPNDEPFVLVQLYKIVPNVLLGOGAKNPNPNVDAHSGVLLQYGMTEMY 420  
 OY 419 YTVLFGVSRLGVLAOLIMSRALGFPLERPKSMSTEGLMKFVDSK 459  
 DB 421 YTVLFGVSRLGVLAOLIMSRALGFPLERPKSMSTEGLMKFVDSK 461

## RESULT 3

739028  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #extl\_change 21-Jan-2000  
 C:Species: Schizosaccharomyces pombe  
 R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, February 1996  
 A:Accession: T39028  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-473 <DEV>  
 A:Cross-references: EMBL:269731; PIDN:CA93617.2; GSPDB:GN00066; SPDB:SPAC6C3.04  
 A:Experimental source: strain 972h-; cosmid c6C3  
 C:Genetics:  
 A:Gene: SPDB:SPAC6C3.04  
 A:Map position: 1  
 A:Genome: nuclear  
 C:Superfamily: citrate (si)-synthase  
 C:Keywords: mitochondrion

Query Match 62.4%; Score 1519; DB 2; Length 473;

Best Local Similarity 61.3%; Pred. No. 5.6e-107; Mismatches 96; Indels 10; Gaps 3;

OY 6 AARLLGTRKNASCLVLAARH-----SASSTNLDIADLIPKBOARKITRROOCHT 57  
 DB 7 ASTRLLASLSLSOASLSRDLNPLFTSSYSTSSSLDLRALIPKBOARKITRROOCHT 66  
 OY 58 TVGQITVDAMYGMRGKGLVETSVLDPDEGIRFRGFSIPBECOKLPRKAGGEEPLPE 117  
 DB 67 DVIGEVITNOMTGARGVRSILWEGSVLDPBGRITFRGTYITPECOKLLPSPNKKOPLPE 126  
 OY 118 GLFWLLVTGHIPTBEOVSWLSKEMAKRAALPSHVYTMNDNPTNLHPMSQLSAAVTALNS 177  
 DB 127 SLFWLLVTGHIPTBEOVSWLSKEMAKRAALPSHVYTMNDNPTNLHPMSQLSAAVTALNS 186  
 OY 178 ESNFARAYAGISRTKYVELLYEDSMOILAPCAAKIYRNLYREGSGTALDINSNDMS 237  
 DB 187 DSAPFAKAYERKMNKRDYKYEEDSMOILAPCAAKIYRNLYREGSGTALDINSNDMS 245  
 OY 238 NFWTMLGTYDHOFTLRLYLTITSDHEGNGVSAHTSHLVGALSADPYLSFAAAMNGL 296  
 DB 246 YNFANVLGFANNEEFVELMRLYLTITSDHEGNGVSAHTSHLVGALSADPYLSFAAAMNGL 305  
 OY 297 AGPLHGLANOEVLWLTOLQKEVGDVSDKLDYIMNTLNSGRVVPYGAHVLKRTDPR 356  
 DB 306 AGPLHGLANOEVLWLTOLQKEVGDVSDKLDYIMNTLNSGRVVPYGAHVLKRTDPR 365  
 OY 357 YTCQREFALKHLPNDEPFVLVQLYKIVPNVLLGOGAKNPNPNVDAHSGVLLQYGMTE 416  
 DB 366 YTCQREFALKHLPNDEPFVLVQLYKIVPNVLLGOGAKNPNPNVDAHSGVLLQYGMTE 425  
 OY 417 YTVLFGVSRLGVLAOLIMSRALGFPLERPKSMSTEGLMKFVDSK 464  
 DB 426 YTVLFGVSRLGVLAOLIMSRALGFPLERPKSMSTEGLMKFVDSK 473

## RESULT 4

T49379

**THIS PAGE BLANK (USPTO)**